

SISTEMA DE APOIO À DECISÃO PARA GESTÃO DA PÓS-
COLHEITA DO CAFÉ

RAFAEL VARGAS MESQUITA DOS SANTOS

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE

DARCY RIBEIRO

CAMPOS DOS GOYTACAZES - RJ
DEZEMBRO - 2012

SISTEMA DE APOIO À DECISÃO PARA GESTÃO DA PÓS-COLHEITA DO CAFÉ

RAFAEL VARGAS MESQUITA DOS SANTOS

Tese apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Produção Vegetal.

Orientador: Prof. Dr. Henrique Duarte Vieira
Coorientador: Prof. Dr. Flávio Meira Borém

CAMPOS DOS GOYTACAZES - RJ
DEZEMBRO - 2012

SISTEMA DE APOIO À DECISÃO PARA GESTÃO DA PÓS-COLHEITA DO CAFÉ

RAFAEL VARGAS MESQUITA DOS SANTOS

Tese apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Produção Vegetal.

Aprovada em 06 de dezembro de 2012

Comissão Examinadora

Prof. Sílvio de Jesus Freitas (D.Sc., Produção Vegetal) - UENF

Prof. Eder Pedroza Isquierdo (D.Sc., Ciência dos Alimentos) - UFLA

Prof. Flávio Meira Borém (D.Sc., Produção Vegetal) - UFLA

Prof. Henrique Duarte Vieira (D.Sc., Produção Vegetal) - UENF
(Orientador)

Dedico esta tese à minha amada esposa Sabrina, por todo o carinho, apoio, respeito e dedicação.

AGRADECIMENTOS

A Deus, pelas misericórdias renovadas a cada manhã.

Ao professor Henrique Duarte Vieira, pela orientação, apoio e confiança durante todo trabalho;

Ao professor Flávio Meira Borém, pela coorientação desde a época da graduação, disponibilidade e incentivo;

Aos professores Pedro Amorim Berbert e Geraldo de Amaral Gravina pelas valiosas sugestões;

Aos doutores Eder Pedroza Isquierdo e Pedro Damasceno de Oliveira pela grande ajuda com os dados de questionários;

Aos professores das disciplinas cursadas, pelos conhecimentos transmitidos;

A Universidade Estadual do Norte Fluminense Darcy Ribeiro, ao Programa de Pós-Graduação em Produção Vegetal e ao Laboratório de Fitotecnia, pela oportunidade de realização deste curso.

Aos colegas do curso de Pós-graduação que, de maneira geral, sempre foram companheiros e amigos;

Aos meus pais, Wilton Mesquita dos Santos e Roziane Vargas Mesquita dos Santos, pela educação e exemplos de vida;

Especialmente à minha esposa, Sabrina Brugnara Vargas, pelo amor, conselhos e compreensão;

A todos aqueles que estiveram comigo, durante esta jornada, os meus agradecimentos;

SUMÁRIO

1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	3
2.1. Sistemas de apoio à decisão	3
2.1.1. Arquitetura de sistemas de apoio à decisão	5
2.2. Processamento do café	8
2.2.1. Processamento via seca	11
2.2.2. Processamento via úmida	12
2.2.3. Custo de processamento do café	15
2.3. Análise estatística multivariada	16
2.3.1. Análise de agrupamento (AA)	17
2.3.2. Análise de componentes principais (ACP)	21
2.3.3. Análise de fatores (AF)	23
3. MATERIAL E MÉTODOS	24
3.1. Definição dos custos dos processamentos a serem calculados	24
3.2. Identificação das variáveis necessárias para o cálculo dos custos definidos	25
3.3. Implementação do sistema de apoio à decisão “Pós-Café”	26

3.4.	Aplicação de questionários aos produtores de café.....	27
3.5.	Análises estatísticas.....	28
3.5.1.	Análise estatística para validação do sistema de apoio à decisão “Pós-Café”	28
3.5.2.	Análise estatística multivariada em dados de custos da fase de pós-colheita do café	28
3.5.3.	Análise estatística para a estimativa da infraestrutura	29
3.5.4.	Análise estatística para o desenvolvimento de cenários utilizando o sistema de apoio à decisão “Pós-Café”	30
4.	RESULTADOS E DISCUSSÃO.....	31
4.1.	Sistema de apoio à decisão “Pós-Café”	31
4.2.	Análise estatística para validação do sistema de apoio à decisão “Pós-Café”	36
4.3.	Análise estatística multivariada para o custo da fase de pós-colheita do café	39
4.3.1.	Análise de agrupamento para os custos da fase de pós-colheita do café	41
4.3.2.	Análise de fatores para os custos da fase de pós-colheita do café	42
4.3.3.	Análise de componentes principais para os custos da pós-colheita do café	45
4.4.	Análise estatística para a estimativa da infraestrutura	47
4.5.	Análise estatística para o desenvolvimento de cenários utilizando o sistema de apoio à decisão “Pós-Café”	57
5.	RESUMO E CONCLUSÕES	65
	REFERÊNCIAS BIBLIOGRÁFICAS.....	68

RESUMO

SANTOS, Rafael Vargas Mesquita dos, D.Sc., Universidade Estadual do Norte Fluminense Darcy Ribeiro, dezembro de 2012. Sistema de apoio à decisão para gestão da pós-colheita do café. Orientador: Prof. Henrique Duarte Vieira.

A escolha do modo de processamento do café, bem como o dimensionamento da infraestrutura necessária para tal operação são decisivos na rentabilidade da atividade cafeeira e dependerão de diversos fatores. Assim, em decorrência da existência de tantas variáveis, é comum o produtor questionar a viabilidade de determinados tipos de processamentos e até mesmo subestimar ou superestimar aspectos de infraestrutura.

Desse modo, o objetivo principal deste trabalho foi desenvolver um sistema de cálculo do custo da pós-colheita do café, facilitando a tomada de decisão quanto aos processamentos de melhor custo-benefício a serem aplicados ao café nesta fase. Dentre os objetivos específicos, temos: estudo dos principais fatores influenciadores no custo da pós-colheita do café, análise de aspectos de infraestrutura das fazendas na fase de pós-colheita do café e detalhamento das propostas de processamentos alternativos mais rentáveis para duas fazendas.

Quarenta e seis fazendas das regiões do Cerrado, Matas de Minas e Sul de Minas Gerais responderam a um questionário elaborado no sentido de possibilitar estas análises.

Ao empregar o sistema de apoio à decisão desenvolvido ("Pós-Café"), foi constatado que os custos na fase de pós-colheita simulados não diferem

estatisticamente dos custos informados pelos responsáveis das fazendas entrevistadas.

A aplicação das técnicas multivariadas de análise de agrupamento, análise de fatores e análise de componentes principais possibilitou complementaridade na explicação dos principais fatores influenciadores no custo da pós-colheita do café.

A partir da utilização do sistema de apoio à decisão para a gestão da pós-colheita do café, também ficou evidenciado que as fazendas entrevistadas, em média, superestimavam a área de terreiro e estimavam corretamente a capacidade dos secadores. Além disso, duas fazendas escolhidas no desenvolvimento de cenários poderiam ampliar sua margem de lucro, caso optassem por processamentos alternativos na fase de pós-colheita do café.

De uma forma geral, o sistema de apoio à decisão para gestão da pós-colheita do café (“Pós-Café”) provou, estatisticamente, ser uma ferramenta confiável, podendo ser utilizada nas mais diversas finalidades dentro do contexto da fase de pós-colheita.

ABSTRACT

SANTOS, Rafael Vargas Mesquita dos, D.Sc., Universidade Estadual do Norte Fluminense Darcy Ribeiro, dezembro de 2012. Development of a decision support system to analyze the cost of post-harvest coffee. Advisor: Prof. Henrique Duarte Vieira.

The choice of method of processing coffee as well as scaling the infrastructure needed for processing are crucial for the profitability of coffee, and will depend on several factors. Thus, due to the existence of many variables is common producer questioned the viability of certain types of processing, as well as underestimate or overestimate aspects of infrastructure.

Thus, the main objective of this study was to develop a system for calculating the cost of post-harvest coffee and thus facilitate the decision making processes regarding the most cost-effective to be applied to coffee at this stage. The specific objectives a study of the major influencing factors in the cost of post-harvest coffee, examined two aspects of infrastructure of the farms in the post-harvest coffee processing and detailed proposals for two alternative more profitable farms.

Forty-six farms in the regions of the Cerrado, Matas de Minas and southern Minas Gerais answered a questionnaire in order to enable these analyzes.

By employing the decision support system, it was found that the costs in the post-harvest simulated costs do not differ statistically informed the heads of the farms surveyed.

The application of multivariate techniques of cluster analysis, factor analysis, principal component analysis and correspondence analysis, possible complementarities in explaining the main factors that influence the cost of post-harvest coffee.

By employing the decision support system for calculating the cost of post-harvest coffee was also found that the farms surveyed, on average, overestimate the area of the yard and correctly estimate the capacity of the dryers. In addition, two farms chosen in the development of scenarios could expand its profit margin, if they chose alternative processes in post-harvest coffee.

In general the decision support system for calculating the cost of post-harvest coffee ("Pós-Café") proved statistically to be a reliable tool, and can be used in various purposes within the context of post-harvest.

1. INTRODUÇÃO

A agricultura brasileira é um setor econômico estratégico para a estabilização da economia, sendo o café um dos seus mais tradicionais produtos, tendo grande influência no processo de industrialização.

Em uma época em que o mercado de cafés especiais está se expandindo cada vez mais e o consumidor ficando mais exigente, elevar a qualidade da bebida é fundamental para atender o mercado e aumentar a renda do cafeicultor brasileiro. E um dos fatores determinantes para isto é a pós-colheita.

As questões agronômicas do café brasileiro já estão, de certa forma, bem evoluídas, todavia, no momento em que o café é colhido e encaminhado para o terreiro surgem as principais dúvidas.

Para os produtores de café, consultores e cooperativas ligadas à área de pós-colheita do café, ainda é muito difícil decidir sobre quais processamentos devem ser aplicados ao café após a colheita. Muito deve ser considerado para que decisões como essas possam ser tomadas com maior facilidade por parte dos produtores, consultores e cooperativas.

Diante do exposto, o presente trabalho teve como objetivos: 1. Desenvolver um sistema de apoio à decisão para a gestão da pós-colheita do café; 2. Aplicar técnicas estatísticas multivariadas em dados de custos da fase de pós-colheita do café para identificar os principais fatores influenciadores do custo na fase de pós-colheita do café; 3. Realizar uma análise comparativa da infraestrutura de processamento de quarenta e seis fazendas, utilizando o sistema de apoio à

decisão para a gestão da pós-colheita do café (“Pós-Café”); 4. Desenvolver cenários de processamentos alternativos mais rentáveis, para duas fazendas representativas, utilizando o sistema de apoio à decisão para gestão da pós-colheita do café (“Pós-Café”).

A tese será apresentada no modelo convencional ou clássico, apresentando os seguintes capítulos: introdução, revisão de literatura, material e métodos, resultados e discussão, além de resumo e conclusões.

2. REVISÃO DE LITERATURA

2.1. Sistemas de apoio à decisão

Sistemas de apoio à decisão (SAD) é uma classe de sistemas de informação ou sistemas baseados em conhecimento. Refere-se a um modelo genérico de tomada de decisão que analisa um grande número de variáveis para que seja possível o posicionamento a uma determinada questão.

Decisão é uma escolha entre as alternativas existentes através de estimativas dos pesos destas alternativas. Apoio à decisão significa auxiliar nesta escolha gerando estas estimativas, a evolução ou comparação e escolha. O termo sistema de apoio à decisão tem sido utilizado de diferentes formas (após a década de 80) e tem recebido diferentes definições de acordo com o ponto de vista de cada autor. Finlay e Forghani (1998) definem o SAD, de um modo geral, como um sistema computacional que auxilia o processo de tomada de decisão. Turban (1995) define, mais especificamente, como um interativo, flexível e adaptável sistema de informação, especialmente desenvolvido para apoiar a solução de um problema gerencial não estruturado no sentido de aperfeiçoar a tomada de decisão. Além de utilizar dados, provê uma interface amigável e permite ao tomador de decisão ter sua própria percepção.

Existe outra definição que se encontra entre estes dois extremos. Para Keen (1980), um SAD concilia os recursos intelectuais individuais com a eficiência dos resultados produzidos pelo computador melhorando a qualidade da decisão

(SAD são sistemas computacionais que apoiam os gerentes tomadores de decisão que são direcionados com problemas semiestruturados). Para Sprague e Watson (1993), SAD's são sistemas computacionais interativos que auxiliam os tomadores de decisão na utilização de dados e modelos solucionados de problemas não estruturados.

Em contraste, Keen (1980) diz que é impossível dar uma definição precisa incluindo todas as facetas do SAD. Segundo o autor, não há definição de sistemas de apoio à decisão, somente de apoio à decisão. Ainda, de acordo com Power (1997), o termo sistema de apoio à decisão é muito desgastado devido a sua utilização para definir muitos tipos de sistemas que dão apoio à tomada de decisão. Como pode-se observar, não há uma definição universal aceita de SAD.

Os SAD's são sistemas computacionais utilizados para resolverem problemas de uma maneira bem parecida com um especialista humano. Eles possuem um conhecimento específico profundo sobre campos restritos do conhecimento.

Existem diversos benefícios associados ao desenvolvimento de um SAD como, por exemplo, distribuição de conhecimento especializado.

Seguem alguns exemplos de sistemas de apoio à decisão, voltados para a área de produção vegetal:

- Sistema de apoio à decisão para utilização no agronegócio: telemetria e tratamento de dados de desempenho de máquina de colheita (Sichonany, 2012);
- Sistema de apoio à decisão para determinação de tarifas em unidades armazenadoras de grãos (Valente et al., 2011);
- Sistema de apoio à decisão para determinação da capacidade de uso dos solos (Cap-Uso, 2012);
- Sistema de apoio à decisão para recomendação do composto de lixo urbano na agricultura (Sirclua, 2012);
- Sistema de apoio à decisão para o planejamento de propriedades agrícolas (Solar, 2012);
- Sistema de apoio à decisão para avaliação da qualidade da água (Acqua-Sist, 2012);

- Sistema de apoio à decisão para diagnóstico de toxi-infecções alimentares (Bac-Sist, 2012);
- Sistema de apoio à decisão para culturas de milho e soja (Narciso, 2012);
- Sistema de apoio à decisão para energização rural (Sisenerg, 2012);
- Sistema de apoio à decisão para identificação e diagnóstico das pragas agrícolas nos campos gerais (Laurindo e Osni Júnior, 2012);
- Sistema de apoio à decisão dedicado à agricultura de precisão: aplicação à cultura da soja (Cruz, 2002);
- Sistema de apoio à decisão na avaliação de risco em investimentos florestais com técnica de simulação discreta (Nobre et al., 2001);

No contexto deste trabalho, tem-se um SAD voltado à área de pós-colheita do café, com o intuito principal de distribuir o conhecimento especializado a produtores, consultores e cooperativas da área.

2.1.1. Arquitetura de sistemas de apoio à decisão

Para Sprague e Watson (1993), muitas são as diversidades existentes entre os sistemas de apoio à decisão encontradas, mas algumas componentes em comum são apresentadas, como: uma interface, um SGBD e um SGBM.

- Interface: interface com o usuário;
- SGBD: sistema gerenciador de banco de dados;
- SGBM: sistema gerenciador de banco de modelos.

A Figura 1 ilustra a arquitetura dos SAD, utilizada no “Pós-Café”.

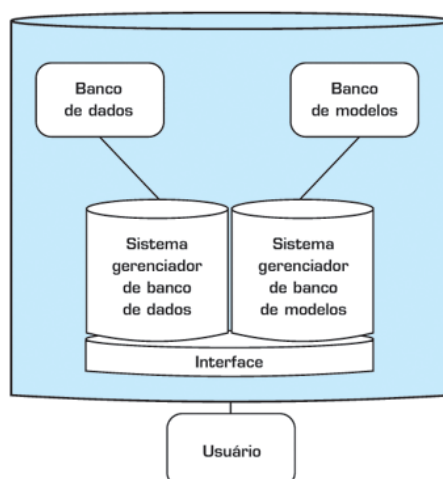


Figura 1. Arquitetura de um SAD (Sprague e Watson, 1993).

A seguir, cada uma dessas componentes será explicada, de acordo com Kock (2003).

Interface

De fundamental importância para o bom funcionamento do homem com a máquina, a interface tem que ser adaptada ao usuário e não o usuário ser adaptado à interface.

A interface do “Pós-Café” é composta por módulo de ajuda, alertas durante a utilização do sistema, gráficos, relatórios resumidos e relatórios completos.

Sistema gerenciador de banco de dados

O SGBD é uma das partes importantes, senão a mais importante, dentro de um SAD, pois sem ele as informações não existiriam e, conseqüentemente, a informática também não, uma vez que esta é ligada totalmente à informação para poder processar algum conteúdo ou resposta às pessoas que vierem a utilizar o sistema.

Os dados estão presentes no SGBD e no BD (banco de dados), sendo que este conjunto, com a adição de um filtro, recebe o nome também de subsistema de dados.

O sistema de gerenciamento de banco de dados escolhido para o desenvolvimento do “Pós-Café” foi o Microsoft Access (Access, 2011), incluído no pacote do Microsoft Office Professional. Ele permite o desenvolvimento rápido de aplicações que envolvem tanto a modelagem e estrutura de dados como também a interface a ser utilizada pelos usuários.

Sistema gerenciador de banco de modelos

Os modelos são gerados por um subsistema de modelos que engloba um SGBM e por um BM constituído de modelos estratégicos, táticos e operacionais.

Turban (1995) classifica os modelos de SAD em sete grupos:

- Enumeração completa: um número finito e pequeno de alternativas é modelado pela análise de decisão. Existem dois tipos distintos em relação aos objetivos: único objetivo e objetivos múltiplos. Tabelas de decisão ou árvores de decisão são duas técnicas utilizadas na análise de único objetivo. Há poucas técnicas em múltiplos objetivos;
- Otimização via algoritmo matemático: programação matemática fornece uma abordagem relativamente imparcial na solução do problema. Programação linear é a técnica mais conhecida utilizada na otimização. O objetivo é encontrar a melhor solução a partir de um grande ou infinito número de alternativas;
- Otimização via fórmula analítica: o objetivo aqui é encontrar a melhor solução usando linhas de fórmulas simples. O modelo de regressão é um exemplo de uma técnica utilizada nesta categoria;
- Simulação: é uma técnica que permite conduzir experiências em modelos digitais da realidade. A simulação é uma das ferramentas mais usadas nos SAD, pois permite descrever situações complexas. O processo de simulação consiste na repetição de uma experiência de modo a obter uma estimativa do efeito produzido por certas ações. A simulação é um

processo descritivo. Não há uma busca da solução ótima para o problema. A simulação descreve ou prevê as características de um dado sistema em diversas circunstâncias. Uma vez calculadas essas características, a melhor entre várias alternativas pode ser escolhida;

- Heurística: o processo de simulação pode ser demorado, complexo e até impreciso. Em tais situações, às vezes, é possível chegar a uma solução satisfatória de forma rápida e menos dispendiosa usando heurísticas. Heurística é o encontro de uma solução "boa o suficiente" para um problema complexo com regras. Como técnicas utilizadas, podemos incluir programação heurística e sistemas especialistas;
- Outros modelos descritivos: outros modelos descritivos envolvem modelos não quantitativos expressos em termos de regras ou fórmulas;
- Modelos prescritivos: o objetivo principal deste tipo de modelo é prever o futuro para um determinado cenário. Técnicas representativas incluem análise de Markov e modelos de previsão.

De acordo com Kock (2003), os SAD podem conter vários modelos: construídos diretamente no software de desenvolvimento do SAD, construídos separadamente e propositadamente, e adicionados pelo SAD quando necessário.

O modelo utilizado no sistema "Pós-Café" foi construído diretamente no software e é baseado em simulação, mais especificamente simulação visual, a qual representa os resultados de uma forma gráfica.

2.2. Processamento do café

O café proveniente da lavoura pode constituir-se de frutos verdes, maduros ou "cereja", supermaduros ou "passa", frutos secos, folhas, ramos, terra, paus e pedras.

Para Borém (2008), a presença de cada um desses constituintes e sua proporção dependerá do sistema e dos cuidados adotados na colheita. Assim, a colheita seletiva resultará na sua totalidade em frutos maduros; colheita por derriça completa, quando realizada precocemente, produzirá cafés com grande

porcentagem de frutos verdes, enquanto colheitas tardias, grande quantidade de frutos secos; nesses dois últimos casos poderão ser produzidos, com maior frequência, cafés com qualidade inferior.

A colheita sobre o pano, mais adequada em relação à colheita no chão, evita o contato dos frutos com o solo e a mistura dos frutos recém-colhidos com os frutos de varrição, na sua maioria já em fase de deterioração. Segundo Theodoro et al. (2002), a qualidade do grão é afetada pelo tipo de colheita, sendo os cafés colhidos no chão, na maioria das vezes, de qualidade inferior ao dos cafés colhidos no pano.

Brando (2004) comenta que os frutos recém-colhidos não deverão ser armazenados em sacos ou silos por períodos superiores a 8 horas, pois, além dos riscos de fermentação, a temperatura poderá atingir valores superiores a 40 °C e assim resultam no defeito ardido. O mesmo autor enfatiza ainda que uma medida emergencial para o armazenamento por períodos mais longos seria a imersão dos frutos frescos em água. Deve-se ressaltar, no entanto, que após a colheita o café deverá seguir imediatamente para o processamento evitando o armazenamento dos frutos úmidos, diminuindo assim os riscos de perda da qualidade, ocorrência de fungos e microtoxinas.

Existem diferenças marcantes na anatomia, na composição química e no teor de água do fruto do cafeeiro em função do seu estágio de maturação. A grande heterogeneidade dos frutos provenientes da lavoura, além de comprometer a qualidade final, irá dificultar as operações posteriores. Quanto mais homogêneo tornar-se o lote colhido ao longo de todas as operações pós-colheita, maior será a eficácia do processamento do café.

A escolha do modo de processamento do café é decisiva na rentabilidade da atividade cafeeira e dependerá de diversos fatores tais como: condições climáticas da região, disponibilidade de capital, tecnologia e equipamentos, exigências do mercado consumidor quanto às características do produto, outorga para uso de água e disponibilidade de tecnologia para o tratamento das águas residuárias. Assim, pode-se dizer que três aspectos são fundamentais na escolha do método de processamento do café: a relação custo/benefício do método de processamento, a necessidade de atendimento à legislação ambiental e o padrão

desejado de qualidade (Saath et al., 2010 ; Resende et al., 2011 ; Borém, 2008 ; Mendonça et al., 2007).

Historicamente, dois diferentes métodos são usados para o processamento do café: a via seca e a via úmida. Na Figura 2 são apresentados estes métodos.

Na via seca, os frutos são processados na sua forma integral, ou seja, com a casca, produzindo frutos secos, conhecidos como café em coco ou café natural. Na via úmida são produzidos os cafés em pergaminho.

No Brasil, Etiópia e Iêmen predomina o processamento por via seca para o café arábica, sendo essa também a via de processamento usada em praticamente todo o café robusta no mundo (Brando, 2004).

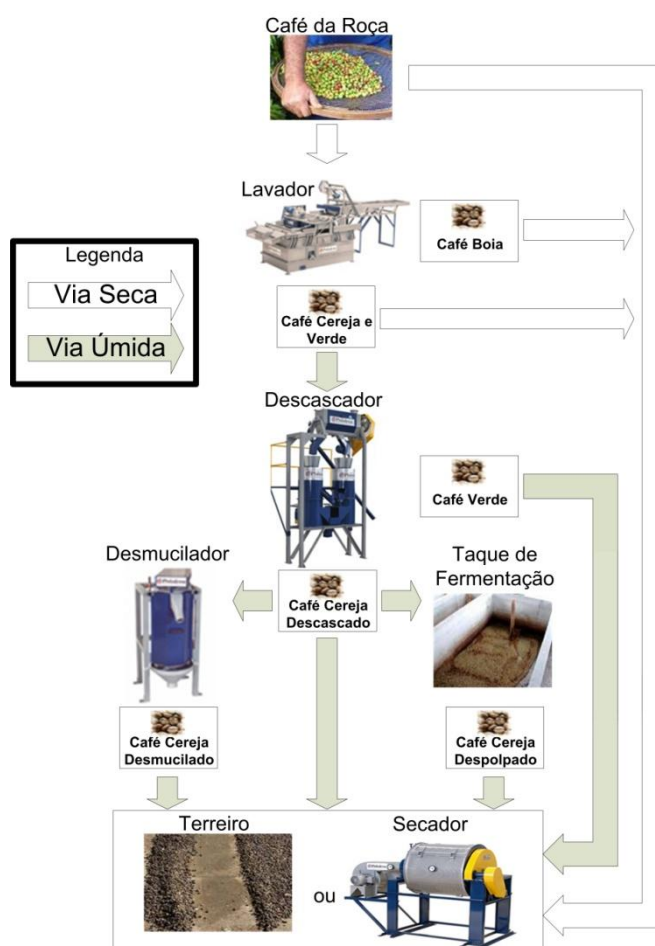


Figura 2. Processamento do café.

2.2.1. Processamento via seca

A produção de café natural, tradicionalmente conhecido como via seca, é o modo mais antigo e mais simples de processar o café e consiste em submetê-lo à secagem na sua forma integral. É largamente usado nas regiões tropicais onde há uma estação seca característica durante o período de colheita.

Segundo Vicent (1987), a via seca consiste na secagem de todos os frutos imediatamente após a colheita. Na verdade, apesar de ser a situação mais frequente, essa é apenas uma das opções desse processo, em geral aplicada por produtores sem a infraestrutura adequada para o processamento. Todo o café, seja ele formado somente por frutos maduros, verdes, secos ou mesmo a mistura desses, quando secados mantendo-se a casca, terá sido processado pela via seca.

Nessa forma de processamento, as impurezas provenientes da lavoura são removidas e os frutos são separados de acordo com sua densidade no lavador, resultando em lotes com diferentes teores de água. Trata-se do processo que menos afeta a condição natural do café, pois todas as suas partes constituintes são mantidas. Além disso, este é o processo que menos agride o meio ambiente, produzindo poucos resíduos sólidos e líquidos, não havendo produção de efluentes com elevado teor de matéria orgânica como acontece na via úmida.

Algumas alterações são observadas durante o processamento via seca conferindo ao café natural características distintas do café despulpado. O exocarpo, originalmente vermelho ou amarelo, torna-se escuro formando o conhecido café coco. O fruto cereja contrai durante a secagem cerca de 40% do seu volume inicial, no entanto, a contração volumétrica de uma mistura de frutos dependerá da proporção de frutos cereja, verde e boia (Borém, 2008).

Segundo alguns autores, parte da mucilagem pode ser transferida para a semente tornando a película prateada mais escura e aderida ao endosperma comparativamente aos cafés produzidos pela via úmida. A transferência, no entanto, de constituintes da mucilagem para a semente ainda não está totalmente comprovada. A coloração dos grãos é marrom-amarelada enquanto a coloração dos grãos despulpados é verde-azulada (Puerta-Quintero, 2000 ; Wilbaux, 1963). Em geral, os cafés produzidos pela via seca apresentam atributos

sensoriais que os distinguem dos cafés produzidos pela via úmida, como menor acidez e maior corpo (Illy & Viani, 1995 ; Puerta-Quintero, 1996 ; Toselo, 1957 ; Vilela, 2002 ; Vincent, 1987 ; Wilboux, 1963), sendo ingrediente essencial para as ligas de café expresso.

A qualidade final do café natural, no entanto, dependerá de diversos fatores desde a colheita até os cuidados durante a secagem. Quando os frutos são colhidos por derricha completa e secados sem qualquer tipo de separação ou cuidado, serão produzidos cafés com bebida dura ou inferior. Em contrapartida, se a colheita for seletiva, secando-se cuidadosamente somente frutos maduros, sem a ocorrência de fermentações, é possível produzir cafés naturais de bebida mole.

As explicações clássicas para a baixa qualidade observada no café processado pela via seca referem-se principalmente à ausência de cuidados na colheita, presença de frutos verdes, frutos brocados e fermentados. Além disso, a secagem mais lenta, pela presença da casca, e o elevado teor de açúcares da mucilagem aumentam os riscos de ocorrerem fermentações, as quais prejudicam a qualidade do café. Essas explicações, no entanto, aparentemente conclusivas, não são suficientes para explicar as diferenças na qualidade do café processado por via seca e via úmida mesmo quando somente frutos maduros são processados com os mesmos cuidados e em condições controladas (Amorim, 1968 ; Angélico et al. 2011 ; Borém, 2008 ; Martins et al., 2005).

2.2.2. Processamento via úmida

Os primeiros relatos do uso do processamento por via úmida na pós-colheita do café datam por volta de 1730, nas Índias Orientais (Tosello, 1957).

Atualmente, a via úmida pode ser conduzida em três formas distintas: removendo-se a casca mecanicamente e a mucilagem por meio de fermentação biológica, resultando no café despulpado, sendo este o método mais tradicional e usado na via úmida, em todo o mundo; removendo-se mecanicamente a casca e parte da mucilagem, resultando no café cereja descascado, também conhecido

como CD, comumente usado no Brasil; ou removendo-se mecanicamente a casca e a mucilagem, produzindo o café desmucilado (Borém, 2008).

Mazzafera (2002) afirma que no Brasil predomina a produção de café pela via seca. A produção de cafés por via úmida (despolpados, descascados e desmucilados) ainda é pequena, comparativamente ao volume total produzido. Ainda assim, sua participação vem crescendo a cada ano, não apenas como necessidade das regiões com maiores limitações para o processamento por via seca, mas como medida para potencializar a obtenção de cafés de bebida fina, mesmo nas regiões consideradas adequadas para a produção de cafés naturais.

Essa melhoria na qualidade a partir do descascamento, no entanto, não é totalmente garantida, pois dependerá de diversos outros fatores como mostram os resultados do trabalho desenvolvido por Silva (2003). Segundo o autor, o levantamento realizado em 32 propriedades da região Sul de Minas Gerais revelou que 75% das amostras apresentaram bebida dura. Apesar de não ter sido encontrada nenhuma amostra com bebida riada ou rio, o autor relata que apenas o descascamento não garantiu uma obtenção exclusivamente de bebidas mole ou estritamente mole.

Descascamento

Independentemente do processo adotado na via úmida, a operação de descascamento é comum para todos eles. O descascamento, ou remoção do exocarpo, é uma operação realizada nos descascadores de cereja.

O descascador de cerejas recebe os cafés cereja e verde após separação dos boias por um lavador. O descascador separa os verdes, retira a polpa das cerejas, separa a polpa e repassa as cerejas não despolpadas, produzindo café pergaminho limpo e sem danos físicos.

Outros equipamentos também são usados para o descascamento dos frutos, como os descascadores de discos, pouco usados no Brasil.

Despolpamento

O despolpamento ocorre logo após o descascamento do café cereja. O café descascado será conduzido para os tanques de fermentação, no qual permanecerá por períodos que variam dentre 12 e 48 horas dependendo da altitude e temperatura do local de processamento.

O objetivo da fermentação é a hidrólise da mucilagem facilitando sua remoção na ação final de lavar o café. Dependendo das condições de secagem, a mucilagem remanescente aderida ao pergaminho pode aumentar os riscos de ocorrerem fermentações indesejáveis, que depreciam a qualidade do produto.

A fermentação pode ocorrer tanto a “seca” como imersa em água. A fermentação mista significa o uso da fermentação a “seco” nas primeiras horas com o objetivo de acidificar o ambiente rapidamente, evitando o desenvolvimento de fungos, seguida pela fermentação com água. Segundo Wootton, citado por Vincent (1987), o café despolpado perde cerca de 1% da matéria seca durante a fermentação e algumas substâncias solúveis são lixiviadas. A perda parcial de alguns compostos pode melhorar a qualidade do café a partir da redução da adstringência e amargor da bebida (Vincent, 1987).

Após a fermentação, o café deverá ser lavado, manual ou mecanicamente, para finalizar a remoção da mucilagem.

Desmucilamento

A operação de desmucilamento mecânico ocorre logo após o descascamento do café com o objetivo de retirar a mucilagem que ainda permanece aderida aos grãos. A remoção é mecânica e ocorre por meio do atrito entre os grãos e dos grãos com um cilindro metálico.

A água é adicionada em pequenas quantidades para lubrificação e limpeza da mucilagem. O café em pergaminho sai pela parte superior da máquina enquanto a água sai pela parte inferior. O consumo de água neste equipamento é variável dependo da regulagem dos registros de água, porém, apesar dos menores volumes de efluente comparando-se com o descascamento, apresenta

maior teor de matéria orgânica, resultando em resíduos líquidos extremamente impactantes ao meio ambiente. Os equipamentos disponíveis no mercado apresentam capacidade variando de 2.000 a 3.600 litros de café por hora.

2.2.3. Custo de processamento do café

Como visto anteriormente, o processamento do café pode ser realizado de formas distintas. A escolha do modo de processamento do café é decisiva na rentabilidade da atividade cafeeira e dependerá de diversos fatores. A relação custo-benefício, no entanto, envolve outras variáveis, tais como: porcentagem de café cereja, porcentagem de café verde, ágio e deságio no momento da comercialização. Assim, em decorrência da existência de tantas variáveis, é comum o produtor questionar a viabilidade do descascamento em relação à produção de cafés naturais. A resposta evidentemente não é simples, necessitando de ferramentas computacionais que facilitem a tomada de decisão. Além disso, o produtor deverá ser capaz de calcular o custo de cada processamento considerando os custos fixos e variáveis (Borém, 2008).

O custo de processamento é composto dos seguintes itens:

- a) custo fixo: definido como sendo todos aqueles que não variam com a quantidade de café processado, portanto sua renovação acontece em longo prazo. Nesse item estão incluídas as depreciações que são definidas com um custo fixo necessário para substituir bens, quando tornados inúteis pelo desgaste físico ou obsolescência. Representa a reserva que o produtor faz durante a vida útil provável dos bens, para a sua posterior substituição;
- b) custo variável: definido como sendo aquele que varia de acordo com a quantidade produzida e cuja duração é igual ou menor que um ciclo produtivo. Nesse item estão incluídos os custos com a mão de obra e energia elétrica;
- c) custo total: o custo total é o parâmetro que melhor avalia o método de processamento e é composto de custo variável e do custo fixo;
- d) custo unitário: o custo unitário é obtido dividindo o custo total pela quantidade de café processado.

2.3. Análise estatística multivariada

Para Kendall (1969), a análise univariada tem por objetivo a análise de uma única variável. Se duas ou mais variáveis forem obtidas, no entanto, essas poderão ser analisadas pela estatística multivariada.

A análise estatística multivariada ou multidimensional se preocupa com as relações entre as variáveis e apresenta duas características: os valores das diferentes variáveis devem ser obtidos sobre os mesmos indivíduos e estas devem ser independentes e consideradas simultaneamente (Kendall, 1969).

A realização de uma análise múltipla sobre uma amostra fornece um melhor entendimento na razão direta do número de variáveis utilizadas e permite considerar, respectivamente, a variabilidade existente nas diversas propriedades avaliadas.

Dentre as técnicas estatísticas de análise multivariada estão à análise de agrupamento, análise de fatores e análise de componentes principais, que são técnicas fundamentalmente diferentes e que podem ser aplicadas independentemente, podendo ser complementares.

A análise multivariada refere-se a todos os métodos analíticos que analisam simultaneamente múltiplas medidas em cada indivíduo ou objeto sob investigação. Qualquer análise simultânea de mais de duas variáveis pode ser considerada análise multivariada.

A análise multivariada conduz aos seguintes objetivos:

- a) Redução de dados ou simplificação estrutural - o fenômeno estudado deve ser representado da maneira mais simples possível, sem sacrificar valiosas informações;
- b) Ordenação e agrupamento - agrupamento de objetos, tratamentos, ou variáveis similares baseados em dados amostrais ou experimentais;
- c) Investigação da dependência entre variáveis - o estudo das relações estruturais entre variáveis muitas vezes é de interesse do pesquisador;
- d) Predição - relações entre variáveis devem ser determinadas para o propósito de predição de uma ou mais variáveis com base na observação de outras variáveis;
- e) Construção e teste de hipóteses.

A seguir, as técnicas multivariadas empregadas para o desenvolvimento deste trabalho serão expostas detalhadamente a fim de que sejam mais bem compreendidas.

2.3.1. Análise de agrupamento (AA)

A análise de agrupamento é uma técnica multivariada que tem por objetivo proporcionar uma ou várias partições na massa de dados, em grupos, por algum critério de classificação, de tal forma que exista homogeneidade dentro e heterogeneidade entre grupos (Mardia et al., 1980).

Conforme Reis (1997), de modo sintético, a técnica pode ser descrita como se segue: dado um conjunto de n indivíduos para os quais existe informação sobre a forma p variáveis, o método de análise de agrupamento procede ao agrupamento dos indivíduos em função da informação existente, de tal modo que os indivíduos pertencentes a um mesmo grupo sejam tão semelhantes quanto possível e sempre mais semelhantes aos elementos do mesmo grupo do que aos elementos dos grupos restantes.

Na análise de agrupamento, é fundamental ter particular cuidado na seleção das variáveis de partida que vão caracterizar cada indivíduo e determinar, em última instância, qual o grupo em que deve ser inscrito. Nesta análise, não existe qualquer tipo de dependência entre as variáveis, isto é, os grupos se configuram por si mesmos sem necessidade de ser definida uma relação causal entre as variáveis utilizadas. Essa análise não faz uso de modelos aleatórios, mas é útil por fornecer um sumário bem justificado de um conjunto de dados. As técnicas são exploratórias e a ideia é, sobretudo, gerar hipóteses, mais do que testá-las, sendo necessária a validação posterior dos resultados encontrados através da aplicação de outros métodos estatísticos (Reis, 1997).

Genericamente, a análise de agrupamento compreende cinco etapas (Aaker et al., 2001):

1. A seleção de indivíduos ou de uma amostra de indivíduos a serem agrupados;

2. A definição de um conjunto de variáveis a partir das quais serão obtidas informações necessárias ao agrupamento dos indivíduos;
3. A definição de uma medida de semelhança ou distância entre os indivíduos;
4. A escolha de um algoritmo de partição/classificação;
5. Por último, a validação dos resultados encontrados.

Dendrograma

Dendrograma é uma representação matemática e ilustrativa de todo o procedimento de agrupamento através de uma estrutura de árvore (Everitt et al., 2001).

Os nós do dendrograma representam agrupamentos. Nós são compostos pelos grupos e ou objetos (grupos formados apenas por ele mesmo) ligados a ele (nó). Se o dendrograma é cortado em um nível de distância desejado, obtém-se uma classificação dos números de grupos existentes nesse nível e dos indivíduos que os formam. O número de grupos dos indivíduos é obtido pelo corte do dendrograma em um nível desejado e então cada componente conectado forma um grupo.

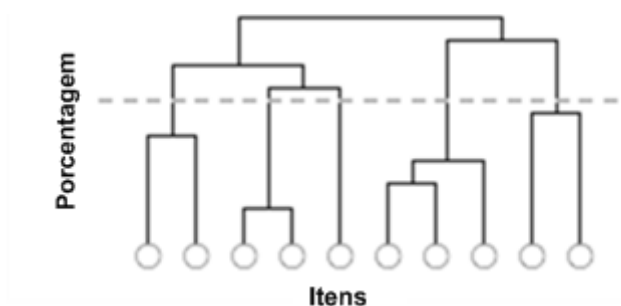


Figura 3. Exemplo de dendrograma

Medidas de distância

Para agrupar indivíduos, é necessária a definição de uma medida de similaridade ou dissimilaridade. Com base nessa medida os indivíduos similares

são agrupados e os demais são colocados em grupos separados (Aaker et al., 2001).

As medidas de dissimilaridade têm papel central nos algoritmos de agrupamentos. Através delas são definidos critérios para avaliar se dois pontos estão próximos, e, portanto, podem fazer parte de um mesmo grupo, ou não.

Existem várias medidas que podem ser utilizadas como medidas de distâncias ou dissimilaridade entre elementos de uma matriz de dados. Cormack (1971) descreve uma série de medidas possíveis: distâncias euclidiana, euclidiana quadrada e euclidiana padronizada, distância corda, distância de Nei, distância absoluta ou City - Block Metric, distância de Minkowski, distância Mahalanobis, distância de Chebychev.

Neste trabalho utilizou-se a distância euclidiana, detalhada abaixo, segundo Cormack (1971).

Distância Euclidiana: a distância entre dois casos (i e j) é a raiz quadrada do somatório dos quadrados das diferenças entre valores de i e j para todas as variáveis ($v = 1, 2, \dots, p$).

$$d_{ij} = \sqrt{\sum_{v=1}^p (X_{iv} - X_{jv})^2}$$

em que:

X_{iv} representa a característica do indivíduo i;

X_{jv} representa a característica do indivíduo j;

p é o número de parcelas na amostra;

v é o número do indivíduo na amostra.

Métodos de agrupamentos

A seguir, são apresentados diversos métodos de agrupamentos que fazem parte dos métodos SAHN (*“Sequencial, Agglomerative, Hierarquic, Nonoverlapping Clustering Methods”*). Vale salientar que, não existe o que se

possa chamar de melhor critério na análise de agrupamentos, mas alguns são mais indicados para determinadas situações do que outros (Kaufmann e Rousseeuw, 1990). É prática comum utilizar vários critérios e fazer a comparação dos resultados. Se tais resultados forem semelhantes, é possível concluir que eles possuem um elevado grau de estabilidade e, portanto, são confiáveis.

Os métodos mais comuns de agrupamento para determinar a distância entre agrupamentos são: ligação simples e ligação completa (Anderberg, 1973).

Método da ligação simples

O método da ligação simples, segundo Orlóci (1978); Gama (1980); e Mardia et al. (1980), é uma técnica de hierarquização aglomerativa e tem como uma de suas características não exigir que o número de agrupamentos seja fixado a priori.

No método da ligação simples (Figura 4), os agrupamentos entre objetos e grupos ou entre grupos são feitos por ligações simples entre pares de objetos, ou seja, a distância entre os grupos é definida como sendo aquela entre os objetos mais parecidos entre esses grupos. Este método leva a grupos longos se comparados aos grupos formados por outros métodos de agrupamentos SAHN (Meyer, 2002). Os dendrogramas, resultantes deste procedimento, são, geralmente, pouco informativos, devido à informação dos indivíduos intermediários que não são evidentes (Carlini-Garcia, 1998).

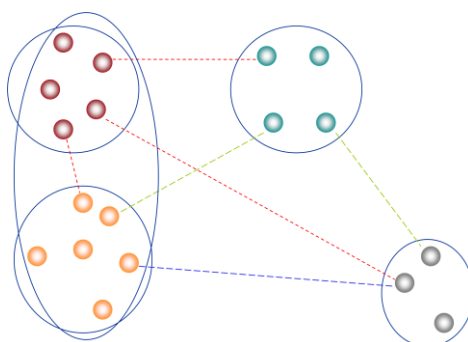


Figura 4. Método da ligação simples (vizinho mais próximo) (Jônatas, 2012)

Método da ligação completa

Este método é também denominado de método do elemento mais distante, sendo uma das técnicas de hierarquização aglomerativa de maior aplicação na análise de agrupamento (Gama, 1980), portanto utilizado neste trabalho. Como no método de ligação simples, aqui também não é exigida a fixação, a priori, do número de agrupamentos.

Conforme Bussab et al. (1990), no método da ligação completa, também denominado vizinho mais distante (Figura 5), a dissimilaridade entre dois grupos é definida como sendo aquela apresentada pelos indivíduos de cada grupo que mais se parecem, ou seja, formam-se todos os pares com um membro de cada grupo e a dissimilaridade entre os grupos é definida pelo par que mais se parece. Este método, geralmente, leva a grupos compactos e discretos, tendo os seus valores de dissimilaridade relativamente grande.

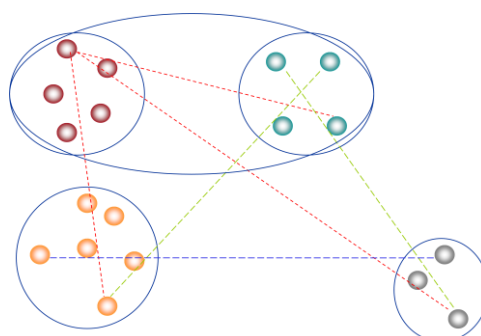


Figura 5. Método da ligação completa (vizinho mais distante) (Jônatas, 2012)

2.3.2. Análise de componentes principais (ACP)

Segundo Pontes (2005), a ACP compreende uma das mais antigas técnicas de análise multivariada que tem por finalidade a análise de estruturas de covariâncias e correlações em matrizes simétricas positivas. Ou seja, essa ferramenta realiza a transformação de variáveis correlacionadas em outro grupo de variáveis não correlacionadas e ainda tem a capacidade de reduzir a

dimensionalidade dos dados. Nesse sentido, a ACP pode ser usada na análise final ou intermediária dos dados.

A análise de agrupamento hierárquico (análise de *cluster*) e de componentes principais são técnicas de estatística multivariada complementares que têm grande aceitação na análise de dados.

É uma técnica que possibilita a análise com um grande número de dados disponíveis e, também, identifica as medidas responsáveis pelas variações relevantes entre os resultados, com as menores perdas de informações possíveis.

A ACP transforma um conjunto de variáveis originais em outro, os componentes principais (CP) com dimensões equivalentes, buscando eliminar algumas das variáveis originais que contenham pouca informação. Cabe ressaltar que a redução de variáveis iniciais somente é possível quando essas não forem independentes e possuem coeficiente de correlação não nulos.

Uma análise de componentes principais tem por objetivo a substituição de um conjunto de variáveis correlacionadas por outro conjunto de novas variáveis não correlacionadas, onde essas são combinações lineares das variáveis originais, dispostas em ordem decrescente de variâncias (Verdinelli, 1980).

Na determinação das CP, calcula-se a matriz de variância-covariância (Σ) ou a matriz de correlação (R), encontram-se os autovalores (Λ) e os autovetores (x) e, finalmente, escrevem-se as combinações lineares, ou seja, as componentes principais, onde cada uma delas é uma combinação linear de todas as variáveis originais (Regazzi, 2001).

A primeira etapa da ACP é a transformação da matriz de dados $n \times p$ em uma matriz quadrada, onde n é o número de indivíduos e p representa um conjunto de variáveis. A interdependência entre as variáveis é representada pela matriz de variância-covariância (Σ) ou a matriz de correlação R.

2.3.3. Análise de fatores (AF)

Conforme Pereira (2006), a análise de fatores (AF) é um processo estatístico que se destina à redução e a sintetização dos dados. Foi desenvolvida por Charles Spearman em 1904.

Malhotra (2011) afirma que a análise de fatores deve ser utilizada nas seguintes situações:

- identificação das dimensões latentes ou fatores que expliquem as correlações entre um conjunto de variáveis;
- identificação de um novo conjunto, menor do que o inicial, de variáveis não correlacionadas em substituição ao conjunto de variáveis correlacionadas na análise multivariada seguinte (regressão ou análise discriminante);
- identificação, em um conjunto maior, de um conjunto menor de variáveis que se destacam para uso em uma análise multivariada posterior.

Cada variável pode ser expressa como uma combinação linear de fatores subjacentes.

Malhotra (2011) explica que a análise de componentes principais (ACP) leva em consideração a variância total dos dados, onde sua maior preocupação é a identificação do número mínimo de fatores que levam a máxima variância nos dados para o emprego na análise que se pretende realizar. Já a análise de fatores (AF) leva em consideração a variância comum nos dados. Sua principal função é a identificação das dimensões subjacente, onde a variância comum é um fator de interesse.

A análise de fatores tem como objetivos principais:

- Analisar as inter-relações entre um grande número de variáveis em termos de poucas, mas não observáveis, variáveis chamadas fatores.
- Verificar se é possível descrever um conjunto de p variáveis em um conjunto menor de índices ou fatores que explicam tanto sobre o fenômeno, quanto o conjunto original.

3. MATERIAL E MÉTODOS

3.1. Definição dos custos dos processamentos a serem calculados

A primeira etapa deste trabalho foi definir quais os tipos de cafés produzidos seriam considerados pelo sistema. Foram eles: café da roça, café boia, café cereja e verde, café verde, café cereja descascado, café cereja desmucilado e café cereja despulpado. Todos estes tipos com a opção de secagem em terreiro ou secador.

Nesta etapa foram definidos os custos a serem calculados para os processamentos do café na fase de pós-colheita.

Na Tabela 1 são apresentados os principais custos calculados pelo sistema, bem como o resumo de suas fórmulas.

Tabela 1. Principais variáveis de custos calculados e suas respectivas fórmulas.

Variável	Fórmula
Custo da Energia Elétrica	$(\text{Total de Energia das Máquinas} * \text{Valor do kWh}) / \text{Quantidade de Sacas}$
Custo da Mão de Obra no Processamento	$(\text{Quantidade de Operadores das Máquinas} * \text{Valor do Dia} * \text{Tempo de Colheita}) / \text{Quantidade de Sacas}$
Custo da Mão de Obra no Terreiro	$(\text{Quantidade de Terreiros da Colheita}$

	Total * Valor do Dia) / Quantidade de Sacas
Custo de Depreciação Total	(Custo de Depreciação de Máquinas + Custo de Depreciação de Terreiro) / Quantidade de Sacas
Custo do Combustível	(Quantidade de Secadores * Tempo Efetivo de Secagem / Tempo Total de Secagem) * Tempo de Colheita * 24 * Custo de Secagem por Hora) / Quantidade de Sacas
(Este custo considera todos os tipos de café submetidos ao secador)	

3.2. Identificação das variáveis necessárias para o cálculo dos custos definidos

Uma vez definidos os principais custos dos processamentos de pós-colheita do café, as variáveis necessárias para o cálculo destes custos foram determinadas.

A partir das fórmulas das variáveis de custos (Tabela 1), definiu-se a necessidade das seguintes variáveis de entrada: produção (sacas), tempo de colheita (dias), porcentagem média de café cereja, porcentagem média de café verde, porcentagem média de café boia, valor da tarifa (R\$/kWh), valor do dia homem (R\$), valor da saca do café padrão (R\$), valor das sacas dos tipos de café (R\$), tipo de terreiro, tanque de fermentação, tratamento de águas residuais, espessura da camada de café no terreiro para cada tipo de café (metros), tempo de secagem no terreiro dos tipos de café (dias), tempo efetivo de secagem no secador dos tipos de café (horas), tempo total de secagem no secador dos tipos de café (horas), tempo máximo de funcionamento das máquinas agrícolas (horas/dia), lavador, peneirão, descascador, ecofiltro, desmucilador, gastos com edificações, mão de obra para o processamento do café, tempo máximo de funcionamento do secador (horas/dia), secador, conjunto elevador, galpão do secador, sistema de aquecimento de ar, combustível utilizado para secar o café.

3.3. Implementação do sistema de apoio à decisão “Pós-Café”

O desenvolvimento da aplicação utilizou as seguintes ferramentas:

Linguagem de Programação

A linguagem de programação escolhida para a implementação foi o Visual Basic.

Os principais motivos para escolha desta linguagem foram:

- Linguagem fácil: considerando a pouca experiência dos desenvolvedores.
- Suporte a vários padrões gráficos: necessário para os resultados das simulações.

Linguagem de Modelagem

Utilizou-se para a modelagem do sistema “Pós-Café” a linguagem UML (*Unified Modeling Language*), que é uma linguagem gráfica para visualização, especificação, construção e documentação de artefatos de sistemas complexos de software (Booch e Rumbaugh, 2006).

A modelagem do sistema foi realizada no programa Rational Rose (Rose, 2011).

Banco de Dados

Foi escolhido o Access (2011) como banco de dados. Este, assim como seu *driver*, não precisa ser instalado na máquina do usuário, tornando simples a instalação do “Pós-Café”.

Gerência de Configuração

A ferramenta utilizada durante o desenvolvimento do projeto, a fim de ajudar na gerência de configuração, foi o Subversion (2011). Esta ferramenta auxilia organizando os arquivos, com a finalidade de manter sempre versões atualizadas em todas as máquinas dos desenvolvedores.

Verificação do Sistema

Durante o desenvolvimento do sistema foram realizados dois tipos de testes:

- Realizados pelos desenvolvedores do sistema: dizem respeito às verificações feitas pelos desenvolvedores a fim de confirmar se os resultados obtidos pelo sistema correspondiam às especificações dos requisitos.
- Realizados por produtores de café: alguns produtores, potenciais auxiliares, receberam cópias do “Pós-Café”, no intuito de que os mesmos retornassem sugestões, melhorias e possíveis alterações no sistema, a fim de que o software pudesse corresponder ao máximo com a realidade da pós-colheita do café.

3.4. Aplicação de questionários aos produtores de café

As 46 fazendas entrevistadas neste trabalho estão localizadas nas regiões: Cerrado, Matas de Minas e Sul de Minas. Possuem altitude média de 938 metros, variando entre 715 metros e 1150 metros, e temperatura média de 24° C, variando entre 15° C e 29° C. Em relação à área utilizada na produção de café declarada pelas mesmas tem-se em média 226,58 hectares.

As regiões estudadas foram escolhidas devido aos diversos tipos de processamentos na fase de pós-colheita do café, possibilitando o estudo mais completo do dimensionamento da infraestrutura nesta fase de processamento.

3.5. Análises estatísticas

3.5.1. Análise estatística para validação do sistema de apoio à decisão “Pós-Café”

Apesar de 46 fazendas terem respondido ao questionário proposto, apenas 34 tiveram condições de declarar a estimativa de custo. Portanto, especificamente para esta análise, os dados de cada questionário respondido por 34 fazendas foram utilizados como entradas em simulações processadas no “Pós-Café”. A proposta foi o estudo detalhado dos relatórios de custo de processamento produzidos pelo sistema.

Para cada fazenda entrevistada realizou-se uma simulação de programa computacional. Estas foram armazenadas em arquivos de computador originados por funcionalidade da própria ferramenta “Pós-Café”, em extensão “.pos”. Dessa forma, as simulações puderam ser analisadas a qualquer momento.

Os dados dos 34 questionários respondidos pelas fazendas, juntamente com as simulações de programa computacional originadas pelo “Pós-Café”, foram comparados com os resultados dos custos declarados de processamento informados pelos responsáveis das fazendas.

Analisando os dados de custos de processamento por simulação e custos declarados informados pelos produtores, e verificando distribuição normal e variância homogênea, foi proposta, para a validação da ferramenta “Pós-Café”, uma análise estatística considerando teste paramétrico, com comparações de médias, especificamente o teste de t de *Student*.

3.5.2. Análise estatística multivariada em dados de custos da fase de pós-colheita do café

Devido ao grande número de variáveis que influenciam os custos de processamentos do café em sua fase de pós-colheita, essa análise tem por objetivo estudar as principais variáveis influentes neste custo, por meio da

utilização de um sistema de apoio à decisão ("Pós-Café") e técnicas estatísticas multivariadas.

Os dados de cada questionário respondido nas 46 fazendas foram utilizados como entrada em simulações processadas no sistema de apoio à decisão para gestão da pós-colheita do café, denominado "Pós-Café" (Borém, 2008).

Os softwares utilizados para realização das análises foram: Statistica e o R. Nesta etapa do trabalho abordaram-se técnicas estatísticas multivariadas para o estudo aprofundado dos dados dos questionários aplicados às 46 fazendas.

Análise de agrupamento das fazendas: foram trabalhadas variáveis quantitativas, por análise de agrupamento, na finalidade de classificá-las em grupos relativamente homogêneos e verificar as similaridades entre as mesmas (Mardia et al., 1997). Os dados foram padronizados e a distância escolhida foi a euclidiana média, devido à ausência de repetições nos tratamentos, conforme sugere Cormack (1971). O método de agrupamento escolhido foi o método da ligação completa, o qual tem tendência a formar grupos compactos (Kaufmann e Rosseeuw, 1990).

Análise de fatores: utilizada para explicar o comportamento de um número relativamente grande de variáveis respostas, em termos de um pequeno número de fatores (Malhotra, 2001). Além disso, também foram estudadas as correlações entre as variáveis, principalmente em relação ao custo simulado da pós-colheita. Para a seleção do número de componentes optou-se pelo critério de Kaiser, citado por Mardia et al. (1997), ou seja, autovalores maiores que 1.

Análise de componentes principais: teve caráter confirmatório, tanto com objetivo de confirmação dos grupos de variáveis formados pela análise de agrupamento, quanto em relação às correlações identificadas na análise de fatores (Jackson, 1980).

3.5.3. Análise estatística para a estimativa da infraestrutura

Para esta análise também foram considerados os dados de cada questionário respondido pelas 46 fazendas.

A proposta foi o estudo detalhado e comparativo entre os dados declarados e simulados, pelo “Pós-Café”, em relação à área de terreiro e a capacidade dos secadores. Sendo realizada uma simulação de programa computacional para cada fazenda entrevistada.

Analisando-se os dados de infraestrutura simulados e declarados pelos produtores nos questionários, e diante de amostras com dimensões superiores a 30 observações, foi proposta, para a análise comparativa da infraestrutura, uma análise estatística considerando teste paramétrico, com comparações de médias, especificamente o teste T-pareado (Souza et al., 2012; Roessler et al., 1978). Para esse procedimento estatístico utilizou-se o programa Statistica, versão 8.

3.5.4. Análise estatística para o desenvolvimento de cenários utilizando o sistema de apoio à decisão “Pós-Café”

Também considerou os dados de questionários respondidos para 46 fazendas.

Realizou-se o desenvolvimento de cenários por intermédio de simulações de custos no “Pós-Café”, as quais consideravam vias de processamento alternativas para cada uma das 46 fazendas.

Analisando-se os dados de margem líquida entre os processamentos originalmente implantados nas fazendas e os processamentos alternativos propostos e diante de amostras com dimensões inferiores a 30 observações, foi proposto inicialmente o teste *Lilliefords* para verificação de normalidade e, uma vez verificada a normalidade dos dados, o procedimento do teste de t de *Student* foi aplicado com o objetivo de comparação de médias com uma constante de referência, no caso zero.

4. RESULTADOS E DISCUSSÃO

4.1. Sistema de apoio à decisão “Pós-Café”

O “Pós-Café” consiste em um sistema de apoio à decisão para gestão da pós-colheita do café, ou seja, emprega o conhecimento humano para resolver problemas que requerem a presença de um especialista nesta fase de processamento do café.

A principal contribuição do “Pós-Café”, a partir da modelagem e sistematização desse conhecimento, é facilitar a tomada de decisão por parte dos produtores, cooperativas e consultores da área, quanto aos processamentos, de melhor custo-benefício, a serem aplicados durante a fase de pós-colheita do café. Esta decisão seria facilitada à medida que se conhecesse o custo dos processos de separação, descascamento, despulpamento, desmucilamento e secagem do café, bem como, a receita originada pela venda dos cafés produzidos após a aplicação destes processamentos.

A ideia básica do sistema é possibilitar que o usuário escolha dois dos possíveis processamentos do café, conforme ilustrado na Figura 2, para que seja realizada uma simulação de custos e comparação entre os mesmos.

Inicialmente, os componentes da Figura 6 estariam desabilitados e, na medida em que o usuário do sistema fosse clicando com o mouse nos tipos de

café, bem como nas máquinas agrícolas, o fluxo de processamento seria montado.

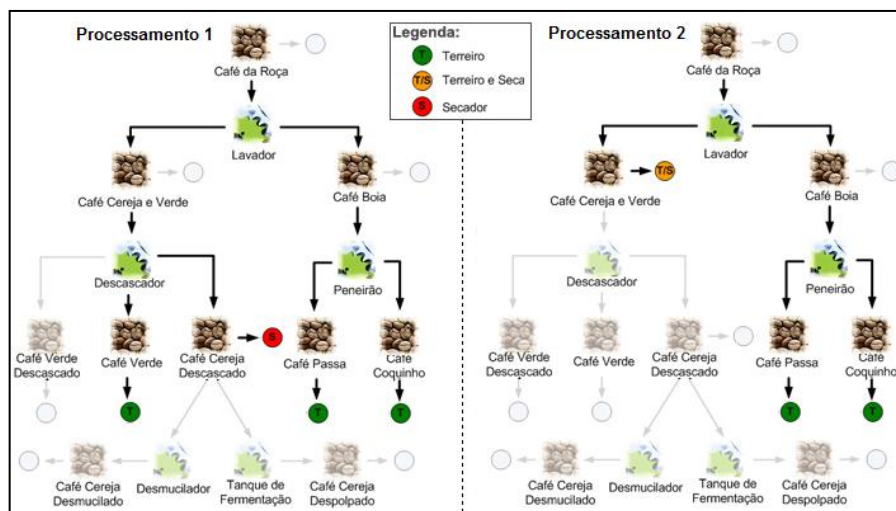


Figura 6. Tela de escolha de processamentos da ferramenta “Pós-Café”.

Após a escolha dos processamentos, o usuário deve fornecer os dados relacionados à produção, de acordo com alguns formulários de entrada de dados (Figura 7).

Entradas do Produtor Produção em Sacas: 7000 Tempo de Colheita (Dias): 135 Café Cereja (%): 30 Café Verde (%): 20 Café Boia (%): 50 Redução no volume do café após descascamento: 40 Quantidade do Café Passa em relação ao Café Boia (%): 20 Descascamento do Café Verde (%): Energia Elétrica Valor da Tarifa (R\$/KWh): 0,29	Trabalhadores Valor do Dia Homem (R\$): 60 Valor da Saca do Café Padrão Café Padrão (R\$): 360 Tipos de Café (R\$) Ágio Deságio Café da Roça: [] [] [] Café Verde: 70 [] [] Café Verde Descascado: [] [] [] Café Boia: [] [] [] Café Cereja e Verde: 0 [] [] Café Cereja Descascado: 20 [] [] Café Cereja Desmucilado: [] [] [] Café Cereja Despolpado: [] [] [] Café Passa: 20 [] [] Café Coquinho: 20 [] []	Selecione o Tipo de Terreiro Tipo de Terreiro: Terreiro de lama asfáltica Custo m²: R\$ [Cadastrar] [Alterar] Tanque de Fermentação Tanques para o processamento do café cereja despolpado. Valor (R\$): [] Vida Útil (Anos): 15 Tratamento de Águas Residuárias Custo (R\$): 30000 Vida Útil (Anos): 20 Volume de Café (Litros) Volume de Café Cereja por Saca: 500 Volume de Café Boia por Saca: 410 Volume de Café Verde por Saca: 650
---	--	---

Figura 7. Tela de entrada de dados da ferramenta “Pós-Café”.

O sistema “Pós-Café” processará e retornará relatórios de receita dos processamentos (Figura 8), de custos dos processamentos (Figura 9) e da margem líquida (Figura 10) entre os mesmos.

Processamento			
Processamento 1			
Processamento 2			
	Total de Sacas Produzidas	Valor da Saca de 60,5 kg	Receita Total
Café da Roça	0,00	0,00	0,00
Café Boia	0,00	0,00	0,00
Café Verde	1.400,00	290,00	406.000,00
Café Verde Descascado	0,00	0,00	0,00
Café Cereja e Verde	0,00	0,00	0,00
Café Cereja Descascado	2.100,00	380,00	798.000,00
Café Cereja Despolpado	0,00	0,00	0,00
Café Cereja Desmucilado	0,00	0,00	0,00
Café Passa	700,00	380,00	266.000,00
Café Coquinho	2.800,00	340,00	952.000,00

Receita Total :

Valor Médio por Saca :

Figura 8. Tela de relatório da receita dos processamentos do “Pós-Café”.

Processamento		Custo (Por Saca de Café)	
Processamento 1		Energia Elétrica	0,77
Processamento 2		Mão de obra Processamento	2,31
Custo Total : <input type="text" value="47.327,98"/>		Mão de obra Terreiro	1,16
Receita Líquida : <input type="text" value="2.374.672,02"/>		Depreciação Total	1,96
		Custo Combustível	0,56
		Custo Total por Saca	6,76

Figura 9. Tela de relatório da receita dos processamentos do “Pós-Café”.

Relatório da Margem Líquida - Comparação entre os Processamentos 1 e 2		
	Processamento 1	Processamento 2
Custo Total/Saca	6,76	10,75
Produção em Sacas	7000	7000
Custo Total	47.327,98	75.237,00
Receita Total	2.422.000,00	2.478.000,00
Receita Líquida	2.374.672,02	2.402.763,00

Margem Líquida
Valor em dinheiro, por saca de café, que o produtor irá ganhar ou perder (custo/benefício) quando o mesmo escolher por um tipo de processamento do café em relação a outro tipo.

Margem Líquida:

Figura 10. Tela de relatório da margem líquida do “Pós-Café”.

A partir dos relatórios informativos, gerados pelo sistema, o produtor terá facilitada a sua tomada de decisão quanto aos processamentos a serem aplicados a sua produção:

- Relatório da receita (Figura 8): Este relatório apresenta o total de sacas produzidas para os tipos de café selecionados no processamento escolhido. Para cada um dos tipos de café também são mostrados os valores em reais de cada saca e a receita total.
- Relatório de custos (Figura 9): Este relatório apresenta os custos por saca, em reais, separados pelas categorias de energia elétrica, mão de obra com o processamento do café (operadores das máquinas agrícolas), mão de obra no terreiro (terreiros trabalhando no revolvimento do café no terreiro), depreciação total (depreciação do terreiro e das máquinas agrícolas) e o custo de combustível (utilizado na secagem mecânica do café).
- Relatório da margem líquida (Figura 10): Este relatório apresenta o valor em dinheiro, por saca de café, que o produtor irá ganhar ou perder quando o mesmo escolher por um tipo de processamento do café em relação a outro tipo. Considerando o exemplo da Figura 10, o produtor perderá quatro reais e um centavo caso opte pelo processamento um (1) ao invés do processamento dois (2).

Todos os relatórios mostrados anteriormente (relatórios de receita líquida, de custos dos processamentos e da margem líquida) levam em consideração os valores de porcentagem de café cereja, porcentagem de café verde, ágio e deságio das sacas de cafés informados nos formulários de entrada de dados (Figura 7).

A Figura 11, todavia, mostra um gráfico de margem líquida gerado pela ferramenta “Pós-Café”.

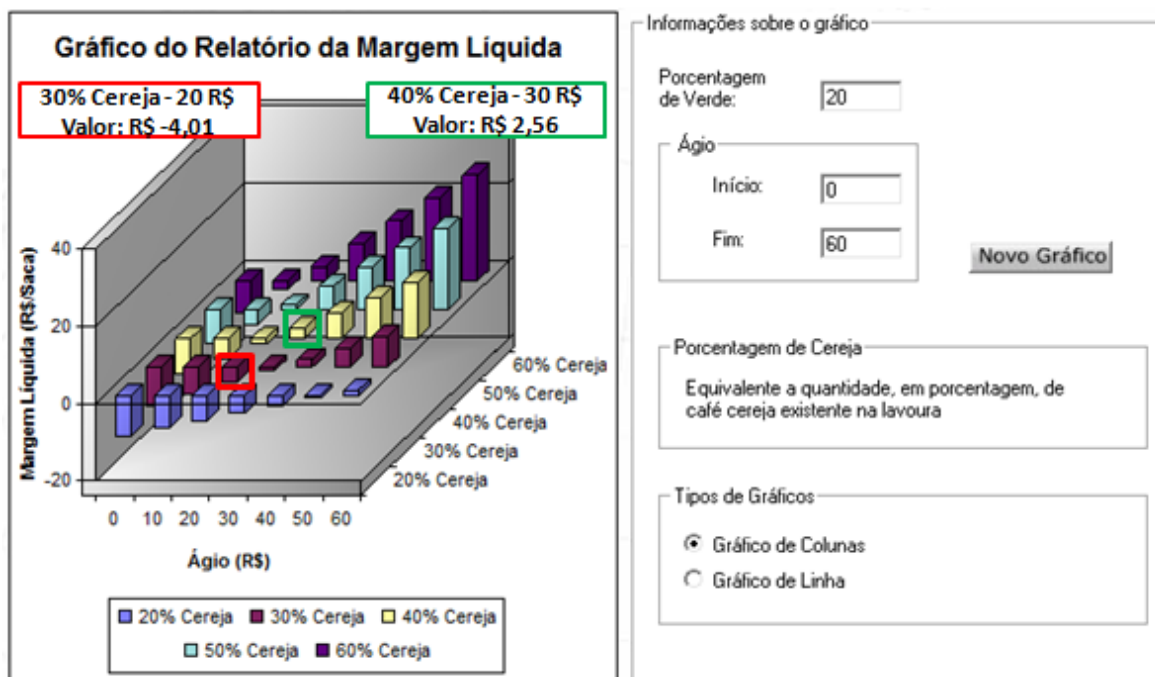


Figura 11. Tela de gráfico da margem líquida da ferramenta “Pós-Café”.

Pelo gráfico da margem líquida o produtor pode visualizar outras situações considerando diferentes ágios (valor em real referente à diferença entre as sacas do tipo de café de maior valor de cada processamento) e porcentagens de cereja (porcentagem de café cereja no momento da colheita).

Conjecturando sobre estes pontos, o produtor pode identificar em que momento é viável ou não, financeiramente, um processamento em relação ao outro.

Pela Figura 11 observa-se destacado o ponto 30% de cereja e 20 reais de ágio, que representa a diferença entre as sacas de maior valor do processamento um (saca do café cereja descascado: R\$380,00) e do processamento dois (saca do café cereja e verde: R\$360,00). Neste ponto, a margem líquida é de aproximadamente menos quatro reais e um centavo (-4,01), ou seja, optar pelo processamento um (1) em relação ao processamento dois (2), para estas condições, fará com que o produtor deixe de ganhar quatro reais e um centavo (-4,01) por saca de café produzido. Para outras condições, por exemplo, o ponto 40% de cereja e 30 reais de ágio, o processamento dois (2) é preferível em

relação ao processamento um (1) no valor de dois reais e cinquenta e seis centavos (2,56).

Resumindo, o gráfico é um recurso da ferramenta que simula novas situações no sentido de auxiliar o produtor na tomada de decisões.

4.2. Análise estatística para validação do sistema de apoio à decisão “Pós-Café”

Para a validação do sistema “Pós-Café” foram comparados os custos declarados por saca de café durante seu processamento na fase de pós-colheita, informados pelos responsáveis das fazendas entrevistadas, com os custos simulados no sistema “Pós-Café”, considerando como dados de entrada as respostas dos questionários aplicados às mesmas fazendas.

A estatística de validação do “Pós-Café”, portanto, considerou os dados de custos declarados e simulados de 34 fazendas, conforme mostra a Tabela 2.

Tabela 2. Dados de custos declarados e simulados das fazendas do sul de Minas Gerais.

Fazenda	Custo Declarado	Custo Simulado
1	R\$ 6,00	R\$ 6,98
2	R\$ 2,10	R\$ 4,08
3	R\$ 9,00	R\$ 8,66
4	R\$ 15,00	R\$ 11,82
5	R\$ 6,00	R\$ 9,36
6	R\$ 15,00	R\$ 7,76
7	R\$ 20,00	R\$ 17,72
8	R\$ 8,50	R\$ 8,04
9	R\$ 10,00	R\$ 11,17
10	R\$ 12,22	R\$ 8,59

11	R\$ 8,45	R\$ 5,24
12	R\$ 8,50	R\$ 8,35
13	R\$ 12,00	R\$ 9,61
14	R\$ 7,15	R\$ 10,00
15	R\$ 6,00	R\$ 8,98
16	R\$ 10,04	R\$ 8,13
17	R\$ 13,02	R\$ 9,32
18	R\$ 10,30	R\$ 10,20
19	R\$ 7,00	R\$ 7,01
20	R\$ 6,00	R\$ 6,95
21	R\$ 4,00	R\$ 8,55
22	R\$ 5,00	R\$ 9,27
23	R\$ 5,80	R\$ 6,05
24	R\$ 2,50	R\$ 4,54
25	R\$ 8,00	R\$ 5,25
26	R\$ 9,50	R\$ 14,16
27	R\$ 9,50	R\$ 9,41
28	R\$ 8,80	R\$ 9,39
29	R\$ 8,00	R\$ 10,85
30	R\$ 8,00	R\$ 6,92
31	R\$ 12,00	R\$ 11,69
32	R\$ 10,00	R\$ 9,25
33	R\$ 10,00	R\$ 9,36
34	R\$ 10,00	R\$ 15,24

Para a validação estatística foi utilizado o teste t de *Student*. O teste t é aplicado para testar hipóteses referentes a médias populacionais, quando as variáveis apresentam-se normalmente distribuídas com variâncias desconhecidas.

De acordo com Ribeiro Júnior e Melo (2009), trata-se de um caso de duas populações dependentes, portanto, a variável aleatória de interesse foi a diferença entre os pares de duas amostras, no lugar dos valores delas próprias, que devem ter o mesmo tamanho.

Na análise de estatística descritiva dos dados de custos, temos as seguintes informações, mostradas na Tabela 3.

Tabela 3. Estatística descritiva dos dados de custos de processamento.

Variáveis	Mínimos	Máximos	Perdidos	Válidos
REP	1,0	34,0	0	34
Y1	2,10	20,0	0	34
Y2	4,08	17,7	0	34

Y1 = Custos Declarados de Processamentos

Y2 = Custos Simulados de Processamentos

D = Y1 - Y2 = Diferença entre os Custos Declarados e Simulados de Processamentos

Após a estatística descritiva realizou-se o teste de *Lilliefords* com o objetivo de verificação de normalidade, conforme Tabela 4.

Tabela 4. Verificação de normalidade.

Variáveis	Valor Calculado	Valor (P=0,05)	Valor (P=0,01)
D	0,0746	0,152	0,177

Como o valor calculado (0,0746) é menor do que os valores tabelados ($\alpha = 0,05$ e $\alpha = 0,01$), não se rejeita a hipótese H_0 , portanto, a variável aleatória D segue distribuição normal.

Uma vez verificada a normalidade dos dados, o procedimento do teste de t de *Student* (Tabela 5) foi aplicado com o objetivo de comparação de médias. As hipóteses testadas foram: H_0 ($\mu_D=0$) vs H_a ($\mu_D \neq 0$), em que D representa a diferença entre as duas populações.

Tabela 5. Teste de t de *Student* para comparação de médias.

Variáveis	Dados	Médias	Desvios	T	GL	Probabilidade
Y1		8,9229	3,6222			
D	34	-0,1329	2,8073	-0,276	33	0,7842
Y2		9,0559	2,8611			

Observa-se que a diferença média foi igual à zero ($P=0,7842 > \alpha=0,05$), logo, não existe diferença entre as médias das variáveis Y1 e Y2.

Assim, pode-se afirmar que os resultados de cálculos de custo de processamento na fase de pós-colheita do café pelo sistema “Pós-Café” estão coerente com os custos informados pelos produtores responsáveis pelas 34 fazendas.

4.3. Análise estatística multivariada para o custo da fase de pós-colheita do café

A Tabela 6 apresenta as respostas das 46 fazendas aos questionários aplicados para algumas das variáveis de entrada.

Tabela 6. Dados de 46 fazendas entrevistadas para as variáveis de entrada selecionadas.

Fazenda	QS (sacas)	QDC (dias)	PCVU (%)	CMO (R\$)	TER (m ²)	SEC (litros)	CSIM (R\$)
1	19000	80	30	30	18000	120000	6,98
2	10000	85	0	50	21000	0	4,08
3	6000	120	28,33	42	10000	0	8,66

4	4000	60	0	50	7000	39000	11,82
5	1700	100	0	50	3800	0	9,36
6	2100	60	0	30	5500	0	7,76
7	6000	90	35,83	50	6000	64000	17,72
8	9000	90	11	25	17900	51000	8,04
9	4000	120	50	45	7300	30000	11,17
10	900	45	0	50	3000	0	8,59
11	1500	60	0	40	4000	0	5,24
12	4500	70	0	45	9000	0	8,35
13	5000	80	0	40	8800	39000	9,61
14	6000	75	0	60	10300	0	10
15	13500	60	17	40	21000	60000	8,98
16	2300	70	0	50	9500	0	8,13
17	800	26	0	50	5000	0	9,32
18	8000	50	0	45	6000	0	10,2
19	1300	80	0	35	3000	0	7,01
20	8000	105	0	45	10600	45000	6,95
21	6500	45	0	50	10000	88000	8,55
22	8500	90	0	46	11000	30000	9,27
23	3500	95	0	30	12000	0	6,05
24	4700	80	0	30	10000	0	4,54
25	4000	90	0	40	6000	22000	5,25
26	18000	120	31	40	46600	123000	14,16
27	16000	100	35	40	20000	171000	9,41
28	6000	90	40	55	12000	60000	9,39
29	6000	80	0	40	10000	30000	10,85
30	38000	110	48	30	66000	150000	6,92
31	6000	72	56	45	27000	72000	11,69
32	35000	90	20	45	85000	126000	9,25
33	11000	60	40	45	20000	90000	9,36
34	1800	100	0	50	4500	0	15,24
35	3000	120	78	30	6600	22500	24,23
36	12200	45	65	40	47000	141000	14,77
37	4000	50	57,5	49	10500	65500	25,77
38	16500	140	30	51,5	29500	84000	9,83
39	27500	110	0	26	37300	210000	10,18

40	1804	75	56	40	6500	35000	15,19
41	12000	100	40	30	20000	90000	17,18
42	20000	100	47	42	27700	105000	11,29
43	10000	100	0	42	21400	36000	10,78
44	5312	100	55	75	10500	67500	12,57
45	8000	120	55	30	8000	69000	12,99
46	5000	90	60	30	3000	57500	18,07

As variáveis apresentadas na Tabela 6 possuem os seguintes significados:

- QS - Quantidade de sacas de café colhidas (sacas);
- QDC - Quantidade de dias necessários para a colheita do café (dias);
- PCVU - Porcentagem de café via úmida produzidos (%);
- CMO - Custo de mão de obra por dia de trabalho (R\$);
- TER - Área de terreiro declarada (metros quadrados);
- SEC - Capacidade de secadores declarada (litros);
- CSIM - Custo simulado de pós-colheita por saca de café (R\$).

As seis primeiras variáveis tem como respostas os dados declarados pelos responsáveis das fazendas entrevistadas. Já a variável número 7, a saber, CSIM, possui dados originados por intermédio de simulações de software realizadas na ferramenta “Pós-Café”.

Para os dados referentes aos custos da fase de pós-colheita do café contidos na Tabela 6, foram aplicadas as técnicas de análise multivariada: análise de agrupamento, análise de fatores e análise de componentes principais.

4.3.1. Análise de agrupamento para os custos da fase de pós-colheita do café

Na realização da análise de agrupamento, para dados padronizados, pelo método de ligação completa e distância euclidiana, o resultado foi o dendrograma da Figura 12.

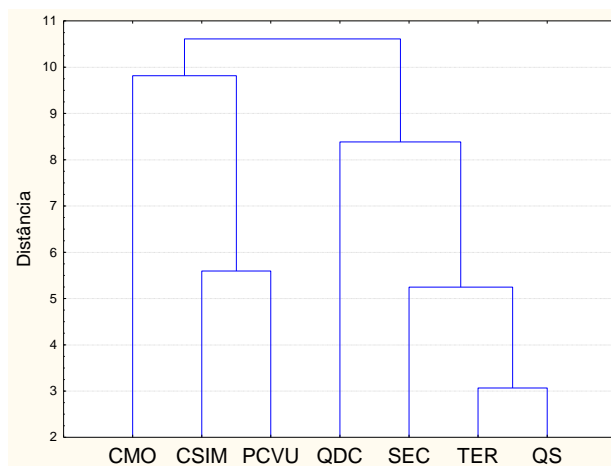


Figura 12. Dendrograma vertical dos dados de custos da fase de pós-colheita do café.

No dendrograma elaborado verifica-se, tomando como base a distância 5,60, a formação de 4 grupos distintos: um grupo composto pela variável CMO; um grupo composto pelas variáveis CSIM e PCVU; um grupo composto pela variável QDC; e um grupo composto pelas variáveis SEC, TER e QS. Uma vez que estes grupos são formados pelas similaridades existentes entre as variáveis, cabe ressaltar que o CSIM e PCVU possuem elevada homogeneidade, ou seja, são variáveis correlacionadas.

4.3.2. Análise de fatores para os custos da fase de pós-colheita do café

Além de se realizar um estudo das correlações entre as variáveis, mediu-se a adequação dos dados através do teste KMO, o qual forneceu um valor de 0,61 e o Bartlett Test com valor de 2669,629 e nível de significância de $p = 0,000$. Através desses valores, demonstra-se que a análise de fatores pode ser conduzida, obtendo um grau de adequação próximo de razoável, conforme a classificação do KMO (Bezerra e Corrar, 2006).

Esta análise explicou o comportamento das sete variáveis respostas, em termos de dois fatores. As variáveis respostas foram agrupadas por meio de suas correlações.

Considerando as sete variáveis e analisando-se os resultados obtidos, pode-se verificar que o 1º e 2º autovalores são maiores que 1 e explicam 65,83% da variância, conforme apresentado pela Tabela 7. Os demais autovalores são menores que 1, portanto, os dados serão resumidos pelo 1º e 2º componentes principais.

Tabela 7. Extração de autovalores: utilizando o método das componentes principais.

Valor	Autovalor	% Variância Total	Autovalor Cumulativo	Cumulativo %
1	2,984130	42,63043	2,984130	42,6304
2	1,624432	23,20617	4,608562	65,8366
3	0,983149	14,04498	5,591711	79,8816
4	0,796313	11,37590	6,388024	91,2575
5	0,286391	4,09130	6,674415	95,3488
6	0,261929	3,74184	6,936344	99,0906
7	0,063656	0,90937	7,000000	100,0000

A Figura 13 mostra os planos fatoriais para custos da fase de pós-colheita do café. Analisando-a, pode-se observar que as variáveis se agrupam por similaridades de explicação, ou seja, estão agrupadas por fatores. As variáveis que explicam ou representam melhor o fator 1 são aquelas que estiverem localizadas mais distantes da origem, em relação ao eixo x. As variáveis que estiverem localizadas próximas à origem do plano fatorial, possuem baixa representatividade. Neste caso, as variáveis que melhor representam o fator 1, num prolongamento do eixo x até a origem dos eixos, são: QS, SEC e TER. Do mesmo modo, as variáveis que melhor representam o fator 2 são: CSIM e PCVU.

As distâncias de cada uma das variáveis analisadas são encontradas traçando-se segmentos de reta perpendiculares ao eixo considerado. Assim, as variáveis que estiverem localizadas mais distantes da origem, a partir da perpendicular traçada, terão maior influência sobre o fator analisado.

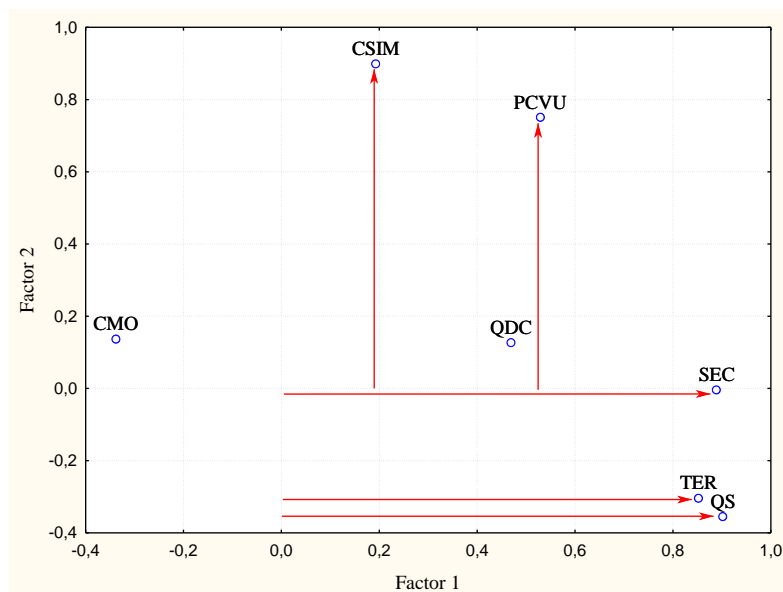


Figura 13. Planos fatoriais para custos da fase de pós-colheita do café.

As correlações entre as variáveis foram determinadas, conforme Tabela 8.

Tabela 8. Correlações para custos da fase de pós-colheita do café.

	QS	QDC	PCVU	CMO	TER	SEC	CSIM
QS	1,00	0,33	0,19	-0,25	0,90	0,79	-0,11
QDC	0,33	1,00	0,28	-0,18	0,22	0,24	0,09
PCVU	0,19	0,28	1,00	-0,07	0,23	0,44	0,65
CMO	-0,25	-0,18	-0,07	1,00	-0,15	-0,22	0,03
TER	0,90	0,22	0,23	-0,15	1,00	0,69	-0,06
SEC	0,79	0,24	0,44	-0,22	0,69	1,00	0,17
CSIM	-0,11	0,09	0,65	0,03	-0,06	0,17	1,00

Analisando-se os resultados da matriz de correlação (Tabela 8) entre as variáveis, observou-se que a maioria dos valores eram inferiores a 0,7 indicando uma correlação fraca entre as variáveis. Apenas entre as variáveis QS e TER; QS e SEC existiam valores de correlações fortes, concluindo-se uma interligação entre as variáveis citadas e que a análise das mesmas deveria ser realizada conjuntamente e não de forma isolada.

É importante, todavia, notar que a variável mais correlacionada (0,65) ao custo simulado (CSIM) foi a porcentagem de café via úmida (PCVU), ou seja, quando a porcentagem de café via úmida aumenta, o valor do custo simulado também tende a aumentar.

4.3.3. Análise de componentes principais para os custos da pós-colheita do café

Para esta análise também foram consideradas apenas as duas primeiras componentes principais, conforme Tabela 7.

O círculo de correlação unitário (Figura 14) foi utilizado como complemento para um maior esclarecimento das análises realizadas.

Analisando-se a Figura 14, observa-se que as variáveis que estiverem mais próximas ao círculo unitário, possuem uma maior contribuição em relação àquelas que estiverem mais afastadas. Também deve ser observado o ângulo formado entre duas variáveis, demonstrando maior ou menor correlação entre as mesmas. Assim, TER e SEC; e CSIM e PCVU contribuem fortemente na análise e possuem correlação significativa, devido ao pequeno ângulo formado entre estas variáveis.

Fazendo-se uma sobreposição da Figura 14 sobre o 1° plano fatorial (Figura 15), identificam-se visualmente quais as variáveis que estão relacionadas com os casos em estudo, ou seja, as fazendas. Como exemplo, podemos citar que para a variável CSIM os casos mais significativos são: 2, 24 (menores valores de custos); 35, 37 (maiores valores de custos).

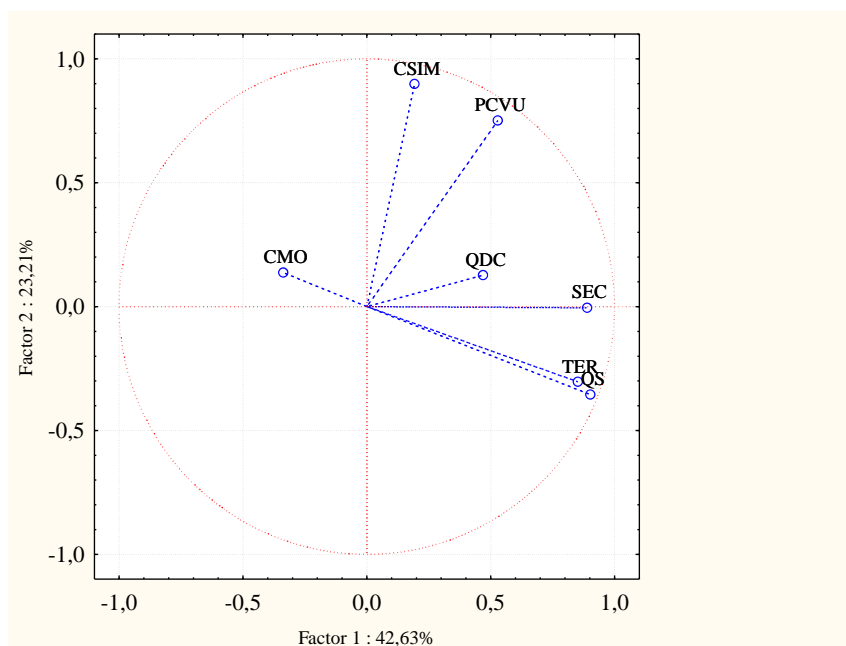


Figura 14. Círculo de correlação unitário para custos da fase de pós-colheita do café.

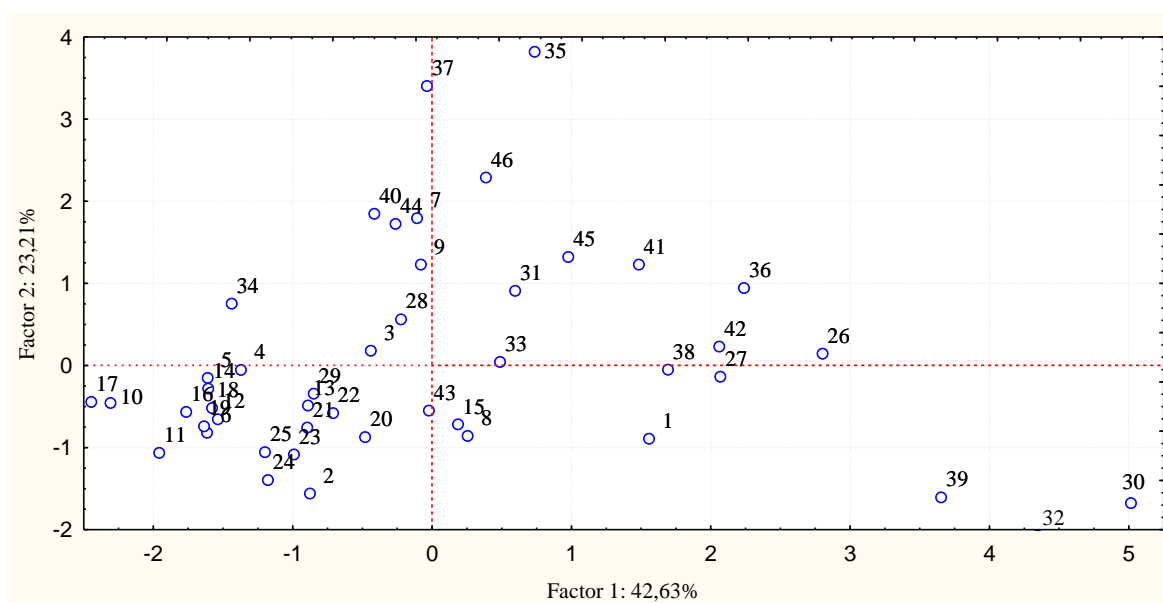


Figura 15. Distribuição da nuvem de pontos por fazendas para custos da fase de pós-colheita.

Analisando-se o gráfico anterior, conclui-se que diversos casos contribuem na formação da combinação linear dos fatores 1 e 2, tornando-se muito difícil a

identificação dos casos mais representativos, podendo-se citar: 30 e 32 para o fator 1; 35 e 37 para o fator 2.

Observa-se uma distribuição bastante uniforme das fazendas, ocorrendo uma maior concentração em torno da origem. Esses casos não apresentam representatividade significativa, em relação aos fatores 1 e 2. Também ocorrem algumas sobreposições, ou seja, algumas fazendas foram influenciadas pelos fatores 1 e 2 de forma muito parecida.

4.4. Análise estatística para a estimativa da infraestrutura

Respostas aos questionários aplicados às 46 fazendas

As informações solicitadas pelo questionário tratam as variáveis de entrada necessárias para o cálculo dos custos de processamentos da fase de pós-colheita. A Tabela 9 apresenta as respostas das 46 fazendas para algumas das variáveis de entrada. Estes valores foram utilizados na simulação da infraestrutura no “Pós-Café”.

Tabela 9. Principais dados, dos questionários aplicados às 46 fazendas, utilizados nas simulações do “Pós-Café”.

ID	N (%)	B (%)	VE (%)	VD (%)	CV (%)	CD (%)	CDM (%)	CDP (%)	SAC (un.)	CO (dias)	Kwh (R\$)	PMO (R\$/Dia)	ESP (m)	TT (dias)	TS (horas)	LAV (un.)	DESC (un.)	DESM (un.)	SEC (un.)
1	15	55	10	0	0	0	20	0	19000	80	0,29	30,00	0,06	8,00	17,00	1	1	1	2
2	100	0	0	0	0	0	0	0	10000	85	0,29	50,00	0,10	6,00	-	0	0	0	0
3	16,7	55	3,33	0	0	25	0	0	6000	120	0,29	42,00	0,03	6,25	-	2	2	0	4
4	20	45	0	0	35	0	0	0	4000	60	0,29	50,00	0,04	5,00	27,00	1	1	0	2
5	100	0	0	0	0	0	0	0	1700	100	0,29	50,00	0,08	10,00	-	0	0	0	0
6	100	0	0	0	0	0	0	0	2100	60	0,29	30,00	0,05	12,00	-	1	0	0	0
7	12,5	51,7	12,5	3,33	0	0	20	0	6000	90	0,29	50,00	0,06	7,00	36,00	1	1	1	4
8	44	45	0	0	0	0	11	0	9000	90	0,29	25,00	0,05	5,33	24,00	1	1	1	2
9	10	40	20	0	0	0	30	0	4000	120	0,29	45,00	0,05	5,75	25,00	1	1	1	2
10	100	0	0	0	0	0	0	0	900	45	0,15	45,00	0,08	9,75	-	1	0	0	0
11	100	0	0	0	0	0	0	0	1500	60	0,15	40,00	0,10	8,50	-	0	0	0	0
12	100	0	0	0	0	0	0	0	4500	70	0,15	45,00	0,08	7,50	-	0	0	0	0
13	10	60	0	0	30	0	0	0	5000	80	0,26	40,00	0,07	7,33	17,50	1	0	0	4
14	100	0	0	0	0	0	0	0	6000	75	0,29	60,00	0,08	9,50	-	1	0	0	0

15	20	63	2,5	0	0	0	14,5	0	13500	60	0,28	40,00	0,05	7,63	38,00	1	1	1	8
16	100	0	0	0	0	0	0	0	2300	70	0,15	50,00	0,06	9,50	-	1	0	0	0
17	100	0	0	0	0	0	0	0	800	26	0,29	50,00	0,07	8,00	-	1	0	0	0
18	100	0	0	0	0	0	0	0	8000	50	0,18	45,00	0,06	8,00	-	0	0	0	0
19	100	0	0	0	0	0	0	0	1300	80	0,29	35,00	0,17	9,25	-	0	0	0	0
20	100	0	0	0	0	0	0	0	8000	105	0,18	45,00	0,06	4,25	20,00	1	0	0	4
21	100	0	0	0	0	0	0	0	6500	45	0,15	50,00	0,04	8,75	25,00	0	0	0	1
22	100	0	0	0	0	0	0	0	8500	90	0,29	46,00	0,06	1,50	24,00	0	0	0	5
23	100	0	0	0	0	0	0	0	3500	95	0,27	30,00	0,04	6,00	-	1	0	0	1
24	100	0	0	0	0	0	0	0	4700	80	0,29	30,00	0,10	9,00	-	1	0	0	0
25	100	0	0	0	0	0	0	0	4000	90	0,17	40,00	0,13	5,50	12,00	0	0	0	1
26	14	55	14	0	0	0	17	0	18000	120	0,29	40,00	0,06	4,00	66,50	2	1	2	9
27	0	65	5	0	0	0	30	0	16000	100	0,15	40,00	0,06	5,50	40,00	2	2	1	8
28	15	45	2	3	0	0	35	0	6000	90	0,29	50,00	0,05	4,40	12,00	1	1	1	3
29	100	0	0	0	0	0	0	0	6000	80	0,29	40,00	0,07	6,00	24,00	0	0	0	2
30	20	32	12	0	0	0	36	0	38000	110	0,16	30,00	0,02	4,50	17,00	3	5	3	10
31	27	17	6	0	0	0	50	0	6000	72	0,17	45,00	0,06	9,17	12,50	2	2	2	4

32	45	35	5	0	0	0	15	0	35000	90	0,19	45,00	0,05	9,40	7,67	3	1	1	9
33	15	45	5	0	0	0	35	0	11000	60	0,29	45,00	0,10	8,25	19,50	1	1	2	6
34	100	0	0	0	0	0	0	0	1800	100	0,29	50,00	0,07	8,50	-	1	0	0	0
35	5	17	8	10	0	0	60	0	3000	120	0,29	30,00	0,04	16,20	12,00	1	3	1	3
36	0	35	15	15	0	0	0	35	12200	45	0,29	40,00	0,06	8,50	36,00	3	5	-	12
37	10	32,5	12,5	7,5	0	0	37,5	0	4000	50	0,29	49,00	0,03	3,20	63,00	2	3	1	8
38	23	43	4	0	0	26	0	0	16500	140	0,29	51,50	0,07	6,00	36,00	1	2	1	5
39	0	50	20	0	0	0	30	0	27500	110	0,29	26,00	0,06	5,70	55,19	1	3	1	10
40	4	40	6	10	0	0	40	0	1804	75	0,29	40,00	0,06	7,75	25,00	1	1	1	2
41	25	35	5	5	0	30	0	0	12000	100	0,25	30,00	0,05	4,80	51,00	1	2	0	6
42	23	30	8	3	0	0	36	0	20000	100	0,19	42,00	0,06	5,00	33,00	2	3	1	6
43	0	50	10	10	0	30	0	0	10000	100	0,29	42,00	0,05	5,00	33,00	2	2	0	4
44	7	38	3	3	0	0	49	0	5312	100	0,29	75,00	0,04	2,50	40,00	1	2	1	5
45	15	30	12,5	12,5	0	0	30	0	8000	120	0,29	30,00	0,03	5,80	25,00	1	2	1	6
46	3	37	10	10	0	0	40	0	5000	90	0,29	30,00	0,10	7,90	52,38	1	2	1	5

Legenda referente às variáveis da Tabela 9:

ID: número identificador da fazenda entrevistada

N: porcentagem de café natural produzido

B: porcentagem de café boia produzido

VE: porcentagem de café verde produzido

VD: porcentagem de café verde descascado produzido

CV: porcentagem de café cereja e verde produzido

CD: porcentagem de café cereja descascado produzido

CDM: porcentagem de café cereja desmucilado produzido

SAC: quantidade de sacas de café produzidas

CO: quantidade de dias de duração da colheita

kWh: preço do quilowatt-hora

PMO: preço da mão de obra por dia de trabalho

MO: mão de obra utilizada no processamento do café (exceto terreiros)

ESP: espessura média do café no terreiro (considerando todos os tipos de café produzidos)

TT: tempo total médio de secagem do café no terreiro (considerando todos os tipos de café produzidos)

TS: tempo total médio de secagem do café no secador (considerando todos os tipos de café produzidos)

LAV: quantidade de lavadores utilizados na fase de pós-colheita do café

DESC: quantidade de descascadores utilizados na fase de pós-colheita do café

DESM: quantidade de desmuciladores utilizados na fase de pós-colheita do café

SEC: quantidade de secadores utilizados na fase de pós-colheita do café

Simulações de infraestrutura para a fase de pós-colheita do café

A Tabela 10 apresenta os resultados das simulações, em relação à infraestrutura, realizadas no sistema “Pós-Café”, para as 46 fazendas.

Tabela 10. Detalhamento da infraestrutura declarada versus simulada para as 46 fazendas entrevistadas.

ID	Área de terreiro declarada (metros)	Área de terreiro simulada (metros)	Capacidade dos secadores declarada (litros)	Capacidade dos secadores simulada (litros)
1	18000,00	15335,38	120000,00	109855,29
2	21000,00	8126,05	0,00	0,00
3	10000,00	1517,05	0,00	0,00
4	7000,00	2170,00	39000,00	35401,67
5	3800,00	954,83	0,00	0,00
6	5500,00	2522,00	0,00	0,00
7	6000,00	3360,94	64000,00	21672,00
8	17900,00	5096,95	51000,00	41288,63
9	7300,00	1191,96	30000,00	10309,00
10	3000,00	1043,00	0,00	0,00
11	4000,00	1257,00	0,00	0,00
12	9000,00	5599,28	0,00	0,00
13	8800,00	2382,81	39000,00	33520,31
14	10300,00	6145,33	0,00	0,00
15	21000,00	6940,58	60000,00	117460,13
16	9500,00	2294,52	0,00	0,00
17	5000,00	1821,97	0,00	0,00
18	6000,00	16796,00	0,00	0,00
19	3000,00	724,47	0,00	0,00
20	10600,00	2380,40	45000,00	31810,91
21	10000,00	6088,33	88000,00	87900,31
22	11000,00	2061,25	30000,00	68708,33
23	12000,00	2315,52	0,00	0,00

24	10000,00	4039,38	0,00	0,00
25	6000,00	782,22	22000,00	18773,33
26	46600,00	3546,96	123000,00	129964,50
27	20000,00	3726,63	171000,00	156600,00
28	12000,00	2528,59	60000,00	13652,00
29	10000,00	3828,12	30000,00	34453,13
30	66000,00	18808,38	150000,00	103174,33
31	27000,00	1557,29	72000,00	65700,00
32	85000,00	24379,70	126000,00	44566,67
33	20000,00	6553,73	90000,00	89694,00
34	4500,00	3108,75	0,00	0,00
35	6600,00	4761,17	22500,00	3367,64
36	47000,00	4645,76	141000,00	192361,47
37	10500,00	4130,67	65500,00	75606,00
38	29500,00	4679,45	84000,00	52032,81
39	37300,00	8707,5	210000,00	252140,58
40	6500,00	1109,15	35000,00	16207,14
41	20000,00	6094,8	90000,00	199987,20
42	27700,00	7012,1	105000,00	102220,95
43	21400,00	4268,4	36000,00	50329,52
44	10500,00	692,22	67500,00	33245,23
45	8000,00	5487,71	69000,00	28389,00
46	3000,00	2316,11	57500,00	62381,54

Os valores indicados em negrito destacam fazendas em que o valor declarado foi menor que o simulado para determinado aspecto analisado, caracterizando, portanto, um déficit de infraestrutura na fazenda.

Estatística comparativa da infraestrutura

Para a análise comparativa da infraestrutura foram estudados os dados declarados, informados pelos responsáveis das fazendas entrevistadas, em relação aos dados simulados no sistema “Pós-Café”.

A Tabela 11 apresenta dados em relação à subestimação e superestimação da infraestrutura por parte das fazendas estudadas.

Tabela 11. Resumo da comparação entre dados declarados versus simulados para a infraestrutura.

Variável	Fazendas de declarações subestimadas (unidade)	Fazendas de declarações superestimadas (unidade)	Fazendas de declarações subestimadas (%)	Fazendas de declarações superestimadas (%)
Área de terreiro	1	45	2,17	97,83
Capacidade de secadores	10	21	21,74	45,65

A maioria das fazendas (97,83%) superestimam dados declarados para o aspecto da fase de pós-colheita mensurado pela variável área de terreiro. Para a capacidade dos secadores esse valor é bem menor, cerca de 45,65% superestimam a capacidade.

Ainda em relação à Tabela 11, uma fazenda tem mau dimensionamento do terreiro, não comportando toda a produção de frutos colhidos e dez fazendas não possuem disponibilidade da capacidade mínima de secadores mecânicos, portanto, estas podem sofrer os malefícios promovidos pelas fermentações que ocorrem nos frutos de café.

Para uma análise mais apurada foi utilizado o teste T-pareado. De acordo com Ribeiro Júnior e Melo (2009), trata-se de um caso de duas populações dependentes, portanto, a variável aleatória de interesse será a diferença entre os pares de duas amostras, no lugar dos valores delas próprias, que devem ter o mesmo tamanho. As hipóteses foram: $H_0 (\mu_{Diferenças}=0)$ vs $H_a (\mu_{Diferenças}\neq 0)$, em que ‘Diferenças’ representa a diferença entre as duas populações. Onde:
 $AT_DEC = \text{Área de terreiro declarada.}$

AT_SIM = Área de terreiro simulada.

AT_D = AT_DEC – AT_SIM = Diferença entre áreas declaradas e simuladas.

CS_DEC = Capacidade dos secadores declarada.

CS_SIM = Capacidade dos secadores simulada.

CS_D = CS_DEC – CS_SIM = Diferença entre a capacidade declarada e simulada.

O procedimento do teste T-pareado (Tabela 12) foi aplicado com o objetivo de comparação de médias. As diferenças entre as duas populações foram testadas para os dois aspectos da fase de pós-colheita: área de terreiro e a capacidade dos secadores.

Tabela 12. Teste T-pareado para comparação de médias.

Variáveis	Dados	Médias	Desvios	T	GL	Probabilidade
AT_D	46	11519,77	13717,26	5,695810	45	0,000001
CS_D	46	2396,226	29755,03	0,546193	45	0,587631

Observa-se que a diferença média foi diferente de zero para a variável AT_D ($P=0,000001 < \alpha=0,05$), portanto, em média, os dados de AT_DEC são 11519,77 maiores que AT_SIM.

A diferença média foi igual à zero para a variável CS_D ($P=0,587631 > \alpha=0,05$), logo, não existe diferença entre as médias das variáveis CS_DEC e CS_SIM.

Para explicação dos dados também foram utilizados os gráficos *Box-plots*. Esses gráficos possibilitam a representação de uma variável aleatória, uma vez que fornece uma ideia de posição, dispersão, assimetria, caudas e dados discrepantes da distribuição. A caixa central (*box*) mostra os dados entre o primeiro e terceiro quartis (*hinge*). Linhas verticais (*whiskers*), iniciando no meio da base (e do topo) da caixa e terminando em valores extremos (denominados adjacentes inferior e superior) indicam, aproximadamente, a variabilidade dos dados (Ferraz et al., 2012 ; McGill et al., 1978).

As Figuras Figura 16 e Figura 17 ilustram *Box-plots* para área de terreiro e capacidade de secadores.

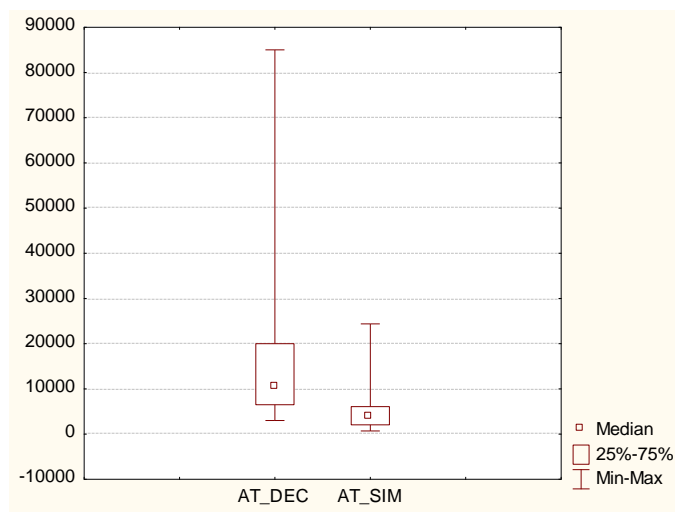


Figura 16. *Box-plots* para dados de área de terreiro.

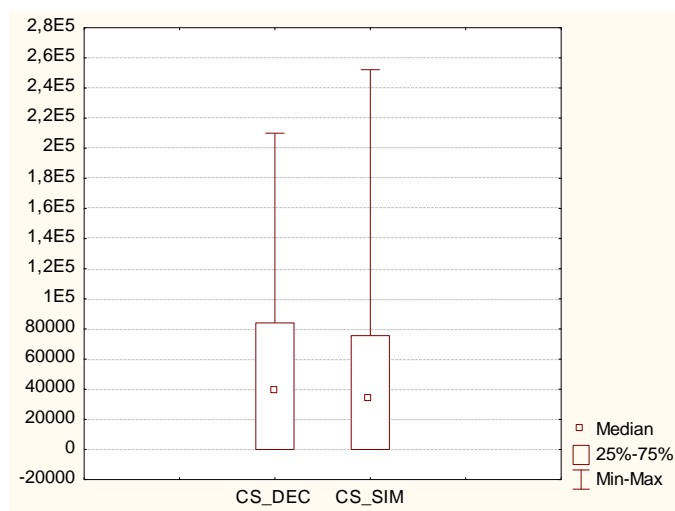


Figura 17. *Box-plots* para dados de capacidade dos secadores.

A partir da análise dos gráficos de *Box-plot* representados, é possível observar que a área de terreiro apresentou valores discrepantes em relação aos dados declarados e simulados. Já os dados de capacidade dos secadores se mostraram bem próximos graficamente para as situações.

A superestimação da área do terreiro, simulada pelo “Pós-Café” e confirmada pelo teste T-pareado, evidencia que os custos de pós-colheita das fazendas com esta superestimativa poderia ser menor. Uma vez que devido à superestimação da área de terreiro, o custo de depreciação do mesmo também está superestimado.

4.5. Análise estatística para o desenvolvimento de cenários utilizando o sistema de apoio à decisão “Pós-Café”

Simulações de programa computacional

Os dados de cada questionário respondido pelas 46 fazendas foram utilizados como entrada em simulações processadas no sistema de apoio à decisão para gestão da pós-colheita do café, denominado “Pós-Café” (Borém, 2008).

As informações solicitadas pelo questionário tratam as variáveis de entrada necessárias para o cálculo dos custos de processamentos da fase de pós-colheita. A Tabela 9 apresenta as respostas das 46 fazendas para algumas das variáveis de entrada. Estes valores foram utilizados na simulação dos custos no “Pós-Café”.

A Tabela 13 apresenta os resultados de custos detalhados das simulações realizadas no sistema “Pós-Café” para os processamentos originais implantados nas 46 fazendas entrevistadas.

Tabela 13. Detalhamentos dos custos simulados no “Pós-Café” para os processamentos originais das 46 fazendas entrevistadas.

Fazenda	CSIM	CEE	CMOP	CMOT	CDT	CC
1	6,98	1,89	0,13	2,27	1,41	1,49
2	4,08	0	0,43	3,4	0,26	0
3	8,66	0,15	2,52	5,04	0,96	0
4	11,82	2,59	2,25	3	2,22	1,77
5	9,36	0	2,94	5,88	0,53	0
6	7,76	0	1,71	5,14	0,91	0
7	17,72	2,41	2,25	6,75	2,71	3,6
8	8,04	1,86	0,75	2,25	1,7	1,48

9	11,17	1,75	2,7	2,7	2,1	2,4
10	8,59	0	2,5	5	1,09	0
11	5,24	0	1,6	3,2	0,44	0
12	8,35	0	0,7	7	0,65	0
13	9,61	1,38	0,64	2,56	1,95	3,07
14	10	0	0,75	8,25	1	0
15	8,98	2,73	0,53	3,2	2,08	0,45
16	8,13	0	1,52	6,09	0,52	0
17	9,32	0	3,25	4,88	1,19	0
18	10,2	0	0,28	7,88	2,05	0
19	7,01	0	2,15	4,31	0,54	0
20	6,95	1,03	0,59	2,36	1,08	1,89
21	8,55	1,11	0,35	3,46	2,5	1,13
22	9,27	3,46	0,49	1,95	1,42	1,96
23	6,05	0	2,44	3,26	0,35	0
24	4,54	0	0,51	3,57	0,45	0
25	5,25	0,6	0,9	1,8	0,93	1,01
26	14,16	4,84	0,8	1,6	1,44	5,47
27	9,41	2,16	0,75	1,75	1,71	3,04
28	9,39	0,77	2,25	2,25	1,63	2,43
29	10,85	2	0,53	3,73	1,88	2,7
30	6,92	1,18	0,26	2,78	1,14	1,56
31	11,69	0,8	1,62	5,94	2,87	0,46
32	9,25	0,51	0,35	5,9	1,02	1,47

33	9,36	1,65	0,74	2,95	2,03	1,99
34	15,24	0	2,78	11,11	1,35	0
35	24,23	0,73	1,2	12	3,82	6,48
36	14,77	6,19	0,15	1,48	3,58	3,39
37	26,03	8,49	0,61	4,9	5,95	6,08
38	9,83	2,7	0,42	3,33	1,09	2,29
39	10,18	4,34	0,1	1,66	1,48	2,59
40	15,19	1,44	1,66	3,33	4,26	4,49
41	17,18	5,14	0,25	2,75	2,74	6,3
42	8,29	2,25	0,21	2,73	1,21	1,89
43	10,78	2,63	0,42	3,36	1,67	2,7
44	12,57	3,98	1,41	1,41	2,38	3,39
45	12,99	2,08	0,45	4,5	2,36	3,6
46	18,07	5,62	0,54	2,7	3,37	5,83

Legenda referente às variáveis da Tabela 13:

Fazenda: número identificador da fazenda entrevistada

CSIM: custo simulado total na fase de pós-colheita (R\$/Saca)

CEE: custo simulado da energia elétrica na fase de pós-colheita (R\$/Saca)

CMOP: custo simulado da mão de obra no processamento na fase de pós-colheita (R\$/Saca)

CMOT: custo simulado da mão de obra no terreiro na fase de pós-colheita (R\$/Saca)

CDT: custo simulado da depreciação total na fase de pós-colheita (R\$/Saca)

CC: custo simulado do gasto com combustível na fase de pós-colheita (R\$/Saca)

Desenvolvimento dos cenários

Nesta etapa o sistema “Pós-Café” foi utilizado para comparar o processamento original de todas as 46 fazendas com as simulações de outros possíveis processamentos aplicáveis às mesmas.

Para as fazendas que apresentavam processamentos originais via seca, foi simulada como alternativa o processamento via úmida com produção de café cereja desmucilado com secagem mecânica.

Já para as fazendas que apresentavam processamentos originais via seca, foi simulada como alternativa o processamento o via seca com produção de café cereja e verde com secagem em terreiro.

A comparação entre o processamento original e o simulado levou em consideração a margem líquida entre os processamentos. A margem líquida é o valor em reais, por saca de café, que o produtor ganha ou perde optando-se por um processamento em relação a outro.

Para esta comparação foram utilizadas cotações obtidas em Carvalhaes (2012). A Tabela 14 mostra os valores de cotações para os diversos tipos de café.

Tabela 14. Cotações dos tipos de café.

Tipo do Café	Preço
Café padrão	R\$ 395,00
Café da roça (natural)	R\$ 340,00
Café verde	R\$ 340,00
Café verde descascado	R\$ 395,00
Café boia	R\$ 360,00
Café cereja e verde	R\$ 395,00
Café cereja descascado	R\$ 440,00
Café cereja desmucilado	R\$ 440,00

A Tabela 15 detalha para cada uma das 46 fazendas: a via de processamento original utilizado (VIA), o custo de pós-colheita para o processamento originalmente utilizado na fazenda (CP1), o custo de pós-colheita

para o processamento alternativo proposto para a fazenda (CP2) e a margem líquida entre os processamentos (ML).

Tabela 15. Margem líquida entre o custo de pós-colheita de processamentos originais e alternativos para as 46 fazendas entrevistadas.

Fazenda	VIA	CP1	CP2	ML
1	ÚMIDA	6,98	5,35	0,75
2	SECA	4,08	5,87	-54,20
3	ÚMIDA	8,66	6,81	7,57
4	SECA	11,82	6,33	-5,94
5	SECA	9,36	13,46	-29,41
6	SECA	7,76	7,83	-30,93
7	ÚMIDA	17,72	12,03	-7,33
8	ÚMIDA	8,04	5,83	-8,01
9	ÚMIDA	11,17	9,14	-0,27
10	SECA	8,59	17,79	-25,23
11	SECA	5,24	11,31	-27,95
12	SECA	8,35	5,99	-45,36
13	SECA	9,61	9,13	-5,95
14	SECA	10,00	7,23	-36,80
15	ÚMIDA	8,98	7,05	2,18
16	SECA	8,13	9,26	-32,89
17	SECA	9,32	16,81	-27,54
18	SECA	10,20	6,66	-37,70
19	SECA	7,01	15,33	-23,69
20	SECA	6,95	7,30	-37,45
21	SECA	8,55	6,72	-28,83
22	SECA	9,27	8,50	-38,78
23	SECA	6,05	8,29	-31,77
24	SECA	4,54	7,98	-28,55
25	SECA	5,25	9,64	-37,72
26	ÚMIDA	14,16	7,49	-1,75
27	ÚMIDA	9,41	5,84	7,18
28	ÚMIDA	9,39	10,04	13,71
29	SECA	10,85	9,25	-34,60
30	ÚMIDA	6,92	7,62	10,28
31	ÚMIDA	11,69	8,70	15,27
32	ÚMIDA	9,25	8,07	2,94
33	ÚMIDA	9,36	8,24	11,88
34	SECA	15,24	15,47	-34,77
35	ÚMIDA	24,23	8,09	9,77

36	ÚMIDA	14,77	7,23	-8,22
37	ÚMIDA	26,03	20,76	6,02
38	ÚMIDA	9,83	8,22	3,15
39	ÚMIDA	10,18	5,61	-2,11
40	ÚMIDA	15,19	9,65	3,66
41	ÚMIDA	17,18	5,65	-0,78
42	ÚMIDA	8,29	6,04	9,76
43	ÚMIDA	10,78	8,26	5,48
44	ÚMIDA	12,57	16,34	24,17
45	ÚMIDA	12,99	7,95	1,53
46	ÚMIDA	18,07	5,53	-0,04

Análise estatística

Para o estudo detalhado dos cenários foi utilizado o teste t de *Student*. O teste t é aplicado para testar hipóteses referentes a médias populacionais, quando as variáveis apresentam-se normalmente distribuídas com variâncias desconhecidas.

Antes da realização desta análise, as fazendas foram agrupadas de acordo com suas vias de processamento.

Na análise de estatística descritiva dos dois grupos temos as seguintes informações, mostradas na Tabela 16 e Tabela 17.

Tabela 16. Estatística descritiva dos dados das 21 fazendas com processamentos originais via seca.

Variáveis	Mínimos	Máximos	Perdidos	Válidos
Fazenda	2,0000	34,0000	0	21
CP1	4,0800	15,2400	0	21
CP2	5,8700	17,7900	0	21
ML	-54,2000	-5,9400	0	21

Tabela 17. Estatística descritiva dos dados das 25 fazendas com processamentos originais via úmida.

Variáveis	Mínimos	Máximos	Perdidos	Válidos
Fazenda	1,00000	46,00000	0	25
CP1	6,92000	26,03000	0	25
CP2	5,35000	20,76000	0	25
ML	-8,22000	24,17000	0	25

Após a estatística descritiva realizou-se o teste de *Lilliefords* com o objetivo de verificação de normalidade para ambos os grupos, conforme Tabela 18 e Tabela 19.

Tabela 18. Verificação de normalidade para os dados de margem líquida (ML) das 21 fazendas com processamentos originais via seca.

Variáveis	Valor Calculado	Valor (P=0,05)	Valor (P=0,01)
ML	0,1767	0,187	0,229

Tabela 19. Verificação de normalidade para os dados de margem líquida (ML) das 25 fazendas com processamentos originais via úmida.

Variáveis	Valor Calculado	Valor (P=0,05)	Valor (P=0,01)
ML	0,0920	0,173	0,220

Como o valor calculado (0,1767 e 0,0920) é menor do que os valores tabelados ($\alpha = 0,05$ e $\alpha = 0,01$) para ambas as vias de processamento, não se rejeitar a hipótese H_0 , portanto, a variável aleatória ML segue distribuição normal tanto na via de processamento via seca, quanto na via úmida.

Uma vez verificada a normalidade dos dados, o procedimento do teste de *t* de *Student* (Tabela 20 e Tabela 21) foi aplicado com o objetivo de comparação de médias com uma constante de referência, no caso zero. As hipóteses testadas

foram: $H_0 (\mu_{ML}=0)$ vs $H_a (\mu_{ML}\neq 0)$, em que ML representa a margem líquida entre os processamentos originais e alternativos das duas vias de processamento.

Tabela 20. Teste de média contra constante de referência para os dados de margem líquida (ML) das 21 fazendas com processamentos originais via seca.

Variável	Dados	Média	Desvios	Constante	T	GL	Probabilidade
ML	21	-31,2410	10,9022	0,00	-13,1317	20	0,000000

Tabela 21. Teste de média contra constante de referência para os dados de margem líquida (ML) das 25 fazendas com processamentos originais via úmida.

Variável	Dados	Média	Desvios	Constante	T	GL	Probabilidade
ML	25	4,2716	7,6242	0,00	2,8013	24	0,009898

No desenvolvimento de cenários para as fazendas com processamento original via seca e processamento alternativo via úmida, com produção de café cereja desmucilado com secagem mecânica, observa-se que a diferença entre a média e a constante de referência foi diferente de zero ($P=0,000000 < \alpha=0,05$), logo, a média da margem líquida neste desenvolvimento de cenários é diferente de zero, a saber, -13,1317. Isso significa que, em média, para as 21 fazendas de processamento via seca analisadas, caso os produtores optassem pelo processamento alternativo via úmida proposto, ganhariam R\$ 13,13 por saca de café produzido.

Já no desenvolvimento de cenários para as fazendas com processamento original via úmida e processamento alternativo via seca, com produção de café cereja e verde com secagem em terreiro, observa-se que a diferença entre a média e a constante de referência foi diferente de zero ($P=0,009898 < \alpha=0,05$), portanto, a média da margem líquida neste desenvolvimento de cenários é diferente de zero, a saber, 2,8013. Isso significa que, em média, para as 25 fazendas de processamento via úmida analisadas, caso os produtores optassem pelo processamento alternativo via seca proposto, perderiam R\$ 2,80 por saca de café produzido.

5. RESUMO E CONCLUSÕES

A escolha do modo de processamento do café, bem como o dimensionamento da infraestrutura necessária para seu processamento são decisivos na rentabilidade da atividade cafeeira e dependerão de diversos fatores. Assim, em decorrência da existência de tantas variáveis, é comum o produtor questionar a viabilidade de determinados tipos de processamentos e até mesmo subestimar ou superestimar aspectos de infraestrutura.

Considerando a argumentação apresentada acima e refletindo sobre a possibilidade do uso potencial de um software que sistematizasse o conhecimento especialista da fase de pós-colheita do café, o presente trabalho foi realizado na expectativa de alcançar os seguintes objetivos:

1. Desenvolver um sistema de apoio à decisão para gestão da pós-colheita do café. Nesse trabalho, investigou-se, também, se o software era estatisticamente confiável para descrever os custos da fase de pós-colheita do café. Caso o resultado fosse positivo, seria então possível utilizar o software para outras análises com base nos dados simulados;

2. Estudar, utilizando técnicas multivariadas, os principais fatores influenciadores no custo da fase de pós-colheita do café;

3. Analisar se a infraestrutura declarada pelos produtores era coerente com a infraestrutura simulada pelo “Pós-Café” necessária para o processamento de secagem do café na fase de pós-colheita;

4. Desenvolver no sistema “Pós-Café” cenários de processamentos alternativos mais rentáveis para duas fazendas com processamentos, respectivamente, via seca e via úmida;

O trabalho experimental necessário para a realização desses objetivos foi então dividido em quatro etapas, permitindo a elaboração de quatro artigos científicos, cada um descrevendo metodologias e resultados específicos. As conclusões desses quatro objetivos encontram-se a seguir.

1. O sistema de apoio à decisão “Pós-Café” provou, estatisticamente, ser uma ferramenta confiável. Os custos de processamento do café em sua fase de pós-colheita calculados pelo “Pós-Café” são compatíveis com os custos declarados.
2. A aplicação das técnicas multivariadas de análise de agrupamento, análise de fatores e análise de componentes principais, possibilitou complementaridade na explicação dos principais fatores influenciadores no custo da pós-colheita do café.

Dentre todas as variáveis estudadas, a de maior correlação com o custo simulado (CSIM) foi a porcentagem de café via úmida produzido (PCVU).

3. Ao empregar o sistema de apoio à decisão para gestão da pós-colheita do café a 46 fazendas, ficou constatado que, em média, pelo teste T-pareado, estas superestimam a área de terreiro e estimam corretamente à capacidade dos secadores.

A superestimação da área do terreiro, simulada pelo “Pós-Café” e confirmada pelo teste T-pareado, evidencia que os custos de pós-colheita das fazendas com esta superestimativa poderia ser menor. Uma vez que devido à superestimação da área de terreiro, o custo de depreciação do mesmo também está superestimado.

4. No desenvolvimento de cenários para as fazendas com processamento original via seca e processamento alternativo via úmida, com produção de café cereja desmucilado com secagem mecânica, observa-se que, caso os produtores optassem pelo processamento alternativo via úmida proposto, ganhariam em média R\$ 13,13 por saca de café produzido.

No desenvolvimento de cenários para as fazendas com processamento original via úmida e processamento alternativo via seca, com produção de café cereja e verde com secagem em terreiro, observa-se que, caso os produtores optassem pelo processamento alternativo via seca proposto, perderiam em média R\$ 2,80 por saca de café produzido.

REFERÊNCIAS BIBLIOGRÁFICAS

- Aaker, D. A., Kumar, V., Day, G. S. (2001) *Pesquisa de marketing*. 1. ed. São Paulo: Atlas, 745 p.
- Access, M. O. (2011) Microsoft Office Access. Disponível em: <<http://office.microsoft.com/en-us/access>>. Acesso em: 8 Junho 2011.
- Acqua-Sist (2012) Sistema de apoio à decisão para avaliação da qualidade da água. Disponível em: <<http://www.ufv.br/ctq/home.htm#acqua-sist>>. Acesso em: 26 Abril 2012.
- Amorim, H. V., Silva, O. M. (1968) Relationship between the polyfenoloxidase activity of coffee beans and quality of the beverage. *Nature*, New York, 219 (5152):381-382.
- Anderberg, M. R. (1973) *Cluster analysis for applications*. 1. ed. New York: Acafenic press, 359 p.
- Angélico, C. L., Pimenta, C. J., Chalfoun, S. M., Chagas, S. J. R., Pereira, M. C., Chalfoun, Y. (2011) Diferentes estádios de maturação e tempos de ensacamento sobre a qualidade do café. *Coffee Science*, Lavras, 6 (1):8-19.
- Bac-Sist (2012) Sistema de apoio à decisão para diagnóstico de toxinfecções alimentares. Disponível em: <<http://www.ufv.br/ctq/home.htm#bacsist>>. Acesso em: 26 Abril 2012.
- Bassab, W. O., Miazaki, E. S., Andrade, D. F. (1990) *Introdução à análise de agrupamentos*. 1. ed. São Paulo: Associação Brasileira de Estatística, 105 p.
- Bezerra, F. A., Corrar, L. J. (2006) Utilização da análise fatorial na identificação dos principais indicadores para avaliação do desempenho financeiro: uma

- aplicação nas empresas de seguros. *Revista Contabilidade & Finanças*, São Paulo, 1 (42):50-62.
- Booch, G., Rumbaugh, J. (2006) *UML, guia do usuário*. 2. ed. Rio de Janeiro: Campus, 463 p.
- Borém, F. M. (2008) Processamento do Café. *In: Borém, F. M. Pós-Colheita do Café*. 1. ed. Lavras: UFLA, Cap. 5, p. 127-158.
- Brando, C. H. J. (2004) Harvesting and Green Coffee Processing. *In: Brando, C. H. J. Coffee: Growing, Processing, Sustainable Production*. 1. ed. Weinheim: Wiley, Cap. 2, p. 605-714.
- Cap-Uso (2012) Sistema de apoio à decisão para determinação da capacidade de uso dos solos. Disponível em: <<http://www.e-science.unicamp.br/sisda/projetos/projetos.php>>. Acesso em: 26 abr. 2012.
- Carlini-Garcia, L. A. (1998) *Estudo da estrutura genética populacional através de marcadores moleculares*. Monografia (Pós-graduação em Agricultura) - Piracicaba - SP, Escola Superior de Agricultura Luiz de Queiroz - ESALQ, 118p.
- Carvalhoes (2012) Escritório Carvalhoes Ltda. Disponível em: <<http://www.carvalhoes.com.br>>. Acesso em: 09 Setembro 2012.
- Cormack, R. (1971) A review of classification. *Journal of the Royal Statistical Society*, 134:321-367.
- Cruz, D. C. (2002) Sistema de apoio à decisão dedicado à agricultura de precisão: aplicação à cultura da soja. Disponível em: <http://www.cartografia.ime.eb.br/trabalhos/MESTRADO/2002_DIONISIO_COSTA_CRUZ_JUNIOR/DissertaCao.pdf>. Acesso em: 26 Abril 2012.
- Everitt, B. S., Landau, S., Leese, M. (2001) *Cluster analysis*. 1. ed. London: Arnold, 207 p.
- Ferraz, G. A. S., Silva, F. M., Costa, P. A. N., Silva, A. C., Carvalho, F. M. (2012) Agricultura de precisão no estudo de atributos químicos do solo e da produtividade de lavoura cafeeira. *Coffee Science*, Lavras, 7 (1):59-67.
- Finlay, P. N., Forghani, M. (1998) A classification of success factors for decision support systems. *Journal of Strategic Information Systems*, 7:53-70.
- Gama, P. M. d. (1980) *Bases da análise de agrupamentos ("Cluster Analysis")*. Dissertação (Mestrado em Estatística e Métodos Quantitativos) - Brasília - DF, Universidade de Brasília - UnB, 229p.
- Illy, A., Viani, R. (1995) *Espresso coffee: the chemistry of quality*. 1. ed. London: Academic, 253 p.

- Jônatas, J. (2012) StatSoft South America Consulting Services StatSoft South America. Disponível em: <<http://www.statsoft.com.br/pt/downloads.php?pg=documentos>>. Acesso em: 01 jun. 2012.
- Kaufmann, L., Rousseeuw, P. J. (1990) *Finding groups in data: an introduction to cluster analysis*. 1. ed. New York: John Wiley, 342 p.
- Keen, P. G. W. (1980) *Decision support systems: A research perspective*. 1. ed. Cambridge: Center for Information Systems Research, 50 p.
- Kendall, M. G. (1969) Discrimination and classification. In: Krishnaian, K. R. *Proc. Int. Symp. on Multivariate Analysis*. 1. ed. New York: Academic Press, p. 165-185.
- Kock, E. (2003) A broader perspective to decision support systems. In: Kock, E. *Decentralising the codification of rules in a decision support expert knowledge base*. 1. ed. Pretoria: University of Pretoria, Cap. 5, p. 90-106.
- Laurindo, A. M., Osni Júnior, T. S. (2012) Sistema de apoio à decisão para identificação e diagnóstico das pragas agrícolas nos campos gerais. Disponível em: <http://www.sbiagro.org.br/pdf/v_congresso/Trabalho89.pdf>. Acesso em: 26 Abril 2012.
- Malhotra, N. K. (2011) *Pesquisa de marketing: uma orientação aplicada*. 6. ed. Porto Alegre: Bookman, 719 p.
- Mardia, A. K. V., T., K. J., Bibby, J. M. (1980) *Multivariate analysis*. 1. ed. London: Academic Press, 521 p.
- Martins, D. R., Camargo, O. A. d., Bataclia, O. C. (2005) Qualidade do grão e da bebida em cafeeiros tratados com lodo de esgoto. *Bragantia*, Campinas, 64 (1):115-126.
- Mazzafera, P. (2002) Degradação de cafeína por microrganismos e o emprego da palha e polpa de café descafeinados na alimentação animal. *Scientia Agricola*, 59 (4):815-821.
- Mcgill, R., Tukey, J. W., Larsen, W. A. (1978) Variations of box plots. *The American Statistician*, New York, 32 (1):12-16.
- Mendonça, L. M. V. L., Pereira, R. G. F. A., Mendes, A. N. G., Borém, F. M., Marques, E. R. (2007) Composição química de grãos crus de cultivares de *Coffea arábica* L. suscetíveis e resistentes à *Hemileia vastatrix* Berg et Br. *Ciência e Agrotecnologia*, Lavras, 31 (2):413-419.
- Meyer, A. S. (2002) *Comparação de coeficientes de similaridade usados em análise de agrupamento com dados de marcadores moleculares dominantes*.

Dissertação (Mestrado em Agronomia) - Piracicaba - SP, Escola Superior de Agricultura Luiz de Queiroz - ESALQ, 106p.

- Narciso, M. G. (2012) Sistema de apoio à decisão para as culturas de milho e soja. Disponível em: <<http://www.academicoo.com/artigo/banco-de-dados-e-sistema-de-apoio-a-decisao-para-as-culturas-de-milho-e-soja>>. Acesso em: 26 Abril 2012.
- Nobre, S. R., Gomes, A. N., Rodriguez, L. C. E. (2001) Sistema de apoio à decisão na avaliação de risco em investimentos florestais utilizando a técnica de simulação discreta estocástica. Disponível em: <<http://www.ipef.br/publicacoes/stecnica/nr34/voluntarios.pdf>>. Acesso em: 26 Abril 2012.
- Orlóci, L. (1978) *Multivariate analysis in vegetational research*. 2. ed. The Hague: Dr. W. Junk B. V. Publishers, 451 p.
- Pereira, A. F. O. A., Carlos Júnior, P., Ramos, E. J. S. (2006) Modelo e análise de previsão de desempenho pela metodologia de análise multivariada de dados: um estudo empírico do setor de energia elétrica. *Revista Contemporânea de Contabilidade*, São Paulo, 3 (5):59-74.
- Pontes, A. C. F. (2005) *Análise de variância multivariada com a utilização de testes não paramétricos*. Tese (Doutorado em Agronomia) - Piracicaba - SP, Escola Superior de Agricultura Luiz de Queiroz, ESALQ, 106p.
- Power, D. J. (1997) What is a DSS? *The On-Line Executive Journal for Data-Intensive Decision Support*, California, 1 (3).
- Puerta-Quintero, G. I. P. (1996) Evaluación de la calidad del café colombiano procesado por via seca. *Cenicafé*, San José: Agronômica Ceres, p. 85-90.
- Puerta-Quintero, G. I. P. (2000) Influencia de los granos de café cosechados verdes, em la calidad física y organoléptica de la bebida. *Cenicafé*, San José: Agronômica Ceres, p. 136-159.
- Regazzi, A. J. (2001) *Análise multivariada*. 1. ed. Viçosa: Universidade Federal de Viçosa, 166 p.
- Reis, E. (1997) *Estatística multivariada aplicada*. 1. ed. Lisboa: Sílabo, 342 p.
- Resende, O., Afonso, P. C., Correa, P. C., Siqueira, V. C. (2011) Qualidade do café conilon submetido à secagem em terreiro híbrido e de concreto. *Ciência e Agrotecnologia*, Lavras, 35 (2):327-335.
- Ribeiro Júnior, I., Melo, A. L. P. (2009) *Guia prático para utilização do SAEG*. 1. ed. Viçosa: Folha Artes Gráficas, 288 p.

- Rose, R. (2011) Rational Rose Enterprise. Disponível em: <www.ibm.com/software/awdtools/developer/rose/>. Acesso em: 8 Junho 2011.
- Saath, R., Borém, F. M., Alves, E., Taveira, J. H. S., Medice, R., Coradi, P. C. (2010) Microscopia Eletrônica de Varredura do Endosperma de Café (*Coffea arabica* L.) durante o processo de secagem. *Ciência e Agrotecnologia*, Lavras, 34 (1):196-203.
- Sichonany, O. R. d. A. O. (2012) Sistema de apoio à decisão para utilização no agronegócio: telemetria e tratamento de dados de desempenho de máquina de colheita. Disponível em: <<http://www.scielo.br/pdf/cr/v41n10/a13411cr5112.pdf>>. Acesso em: 26 Abril 2012.
- Silva, R. F. (2003) *Qualidade do café cereja descascado produzido na região do Sul de Minas Gerais*. Dissertação (Mestrado em Ciência dos Alimentos) - Lavras - MG, Universidade Federal de Lavras - UFLA, 78p.
- Sirclua (2012) Sistema de apoio à decisão para recomendação do composto de lixo urbano na agricultura. Disponível em: <<http://www.e-science.unicamp.br/sisda/projetos/projetos.php>>. Acesso em: 26 Abril 2012.
- Sisenerg (2012) Sistema de apoio à decisão para energização rural. Disponível em: <http://projetos.unioeste.br/campi/nit/index.php?option=com_content&task=view&id=112&Itemid=143>. Acesso em: 26 Abril 2012.
- Solar (2012) Sistema de apoio à decisão para o planejamento de propriedades agrícolas. Disponível em: <<http://www.ufv.br/ctq/home.htm#solar>>. Acesso em: 26 Abril 2012.
- Sprague, J., Watson, H. (1993) *Decision support systems: putting theory into practice*. 3. ed. New Jersey: Prentice Hall College Div, 450 p.
- Subversion, A. (2011) Apache Subversion. Disponível em: <<http://subversion.apache.org>>. Acesso em: 8 Junho 2011.
- Theodoro, V. C. A., Mourão Júnior, M., Guimarães, R. J., Chagas, S. J. R. (2002) Caracterização da qualidade de grãos de cafés (*Coffea arabica* L.) colhidos no pano e no chão, provenientes de sistemas de manejo orgânico, em conversão convencional. *Simpósio de Pesquisa dos Cafés do Brasil, 2*, Vitória: Empresa Brasileira de Pesquisa dos Cafés do Brasil, p. 2358-2366.
- Tosello, A. (1957) Irrigação do Cafezal. *Curso de Cafeicultura, 1*, Campinas: Instituto Agrônomo do Estado de São Paulo, p. 183-193.

- Turban, E. (1995) *Decision Support and Expert Systems*. 4. ed. Califórnia: Prentice-Hall International, 930 p.
- Valente, D. S. M., Queiroz, D. M., Corrêa, P. C., Silva, L. C., Vale, S. M. L. R. d. (2011) Um sistema de apoio à decisão para determinação de tarifas em unidades armazenadoras de grãos. *Engenharia Agrícola*, Jaboticabal, 31 (4):735-744.
- Verdinelli, M. A. (1980) *Análise inercial em ecologia*. Tese (Doutorado em Ciências Biológicas) - São Paulo - SP, Universidade de São Paulo - USP, 162p.
- Vilela, T. C. (2002) *Qualidade de café despolpado, desmucilado, descascado e natural, durante o processo de secagem*. Dissertação (Mestrado em Ciências dos Alimentos) - Lavras - MG, Universidade Federal de Lavras - UFLA, 66p.
- Vincent, J. C. (1987) Green coffee processing. *In*: Clarke, R. J.; Macrae, R. *Technology*. 1. ed. New York: Elsevier, v. 2, Cap. 1, p. 1-199.
- Wilboux, R. (1963) *Agricultural engineering*. 1. ed. Rome: FAO, 293 p.